Against the Dehumanisation of Decision-Making

Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information

by **Guido Noto La Diega***

Abstract: This work presents ten arguments against algorithmic decision-making. These revolve around the concepts of ubiquitous discretionary interpretation, holistic intuition, algorithmic bias, the three black boxes, psychology of conformity, power of sanctions, civilising force of hypocrisy, pluralism, empathy, and technocracy. Nowadays algorithms can decide if one can get a loan, is allowed to cross a border, or must go to prison. Artificial intelligence techniques (natural language processing and machine learning in the first place) enable private and public decision-makers to analyse big data in order to build profiles, which are used to make decisions in an automated way. The lack of transparency of the algorithmic decision-making process does not stem merely from the characteristics of the relevant techniques used, which can make it impossible to access the rationale of the decision. It depends also on the abuse of and overlap between intellectual property rights (the "legal black box"). In the US, nearly half a million patented inventions concern algorithms; more than 67% of the algorithm-related patents were issued over the last ten years and the trend is increasing. To counter the increased monopolisation of algorithms by means of intellectual property rights (with trade secrets leading the way), this paper presents three legal routes that enable citizens to 'open' the algorithms. First, copyright and patent exceptions, as well as trade secrets are discussed. Second, the EU General Data Protection Regulation is critically assessed. In principle, data controllers are not allowed to use algorithms to take decisions that have legal effects on the data subject's life or similarly significantly affect them. However, when they are allowed to do so, the data subject still has the right to obtain human intervention, to express their point of view, as well as to contest the decision. Additionally, the data controller shall provide meaningful information about the logic involved in the algorithmic decision. Third, this paper critically analyses the first known case of a court using the access right under the freedom of information regime to grant an injunction to release the source code of the computer program that implements an algorithm. Only an integrated approach – which takes into account intellectual property, data protection, and freedom of information – may provide the citizen affected by an algorithmic decision of an effective remedy as required by the Charter of Fundamental Rights of the EU and the European Convention on Human Rights.

Keywords: Algorithmic decision-making; algorithmic bias; right not to be subject to an algorithmic decision; GDPR; software copyright exceptions; patent infringement defences; freedom of information request; algorithmic transparency; algorithmic accountability; algorithmic governance; Data Protection Act 2018

Recommended citation: Guido Noto La Diega, Against the Dehumanisation of Decision-Making – Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information, 9 (2018) JIPITEC 3 para 1.

A. Context and scope of the research

1 This work argues that algorithms cannot and should not replace human beings in decision-making, but it takes account of the increase of algorithmic decisions and, accordingly, it presents three European legal routes available to those affected by such decisions.

2 Algorithms have been used in the legal domain for decades, for instance in order to analyse legislation.[1] These processes or sets of rules followed in calculations or other problem-solving operations raised limited concerns when they merely made our lives easier by ensuring that search engines showed us only relevant results.[2] However, nowadays algorithms can decide if one can get a loan,[3] is hired,[4] is allowed to cross a border,[5] or must go to prison.[6] Particularly striking is the episode concerning a young man sentenced in Wisconsin to a six-year imprisonment for merely attempting to flee a traffic officer and operating a vehicle without its owner's consent. The reason for such a harsh sanction was that Compas, an algorithmic risk assessment system, concluded that he was a threat to the community. The proprietary nature of the algorithm did not allow the defendant to challenge the Compas report. The Supreme Court found no violation of the right to due process.[7]

3 Artificial intelligence techniques (natural language processing, machine learning, etc.) and predictive analytics enable private and public decision-makers to extract value from big data[8] and to build profiles, which are used to make decisions in an automated way. The accuracy of the profiles is further enhanced by the linking capabilities of the Internet of Things.[9] These decisions may profoundly affect people's lives in terms of, for instance, discrimination, de-individualisation, information asymmetries, and social segregation.[10]

4 In light of the confusion as to the actual role of algorithms, it is worrying that in "the models of game theory, decision theory, artificial intelligence, and military strategy, the algorithmic rules of rationality replaced the self-critical judgments of reason."[11]

5 One paper[12] concluded by asking whether and how algorithms should be regulated. This work aims to constitute an attempt to answer those questions with a focus on the existing rules on intellectual property, data protection, and freedom of information. In particular, it will be critically assessed whether "the tools currently available to policymakers, legislators, and courts (which) were developed to oversee human decision-makers (...) fail when applied to computers instead."[13]

6 First, the paper presents ten arguments why algorithms cannot and should not replace human decision-makers. After this, three legal routes are presented.[14] The General Data Protection Regulation

* Lecturer in Law (Northumbria University); Director (Ital-IoT Centre for Multidisciplinary Research on the Internet of Things); Fellow (Nexa Center for Internet & Society).

1 William Adam Wilson, 'The Complexity of Statutes' (1974) 37 Mod L Rev 497.

2 The algorithm used by Google to rank search results is covered by a trade secret.

3 More generally, on the use of algorithms to determine the parties' contractual obligations, see Lauren Henry Scholz, 'Algorithmic Contracts' (*SSRN*, 1 October 2016), <https://ssrn.com/abstract=2747701> accessed 1 March 2018.

4 On the negative spirals that automated scoring systems can create, to the point of making people unemployable, see Danielle Keats Citron and Frank Pasquale, 'The scored society: Due process for automated predictions' (2014) 89(1) Washington Law Review 1, 33.

5 Jose Sanchez del Rio et al., 'Automated border control e-gates and facial recognition systems' (2016) 62 Computers & Security 49.

6 As written by Frank Pasquale, 'Secret algorithms threaten the rule of law' (*MIT Technology Review*, 1 June 2017) <https://www.technologyreview.com/s/608011/secret-algorithms-threaten-the-rule-of-law/> accessed 1 March 2018, imprisoning people "because of the inexplicable, unchallengeable judgements of a secret computer program undermines our legal system". For a files $10 million lawsuit related to face-matching technology that allegedly ruined an American man's life see Allee Manning, 'A False Facial Recognition Match Cost This Man Everything' (*Vocativ*, 1 May 2017) <http://www.vocativ.com/418052/false-facial-recognition-cost-denver-steve-talley-everything/> accessed 1 March 2018.

7 *State v Loomis*, 881 N.W.2d 749 (Wis. 2016). Cf Adam Liptak, 'Sent to Prison by a Software Program's Secret Algorithms' (*New York Times*, 1 May 2017), <https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html?_r=0> accessed 1 March 2018.

8 In analysing the algorithms used by social networks, Yoan Hermstrüwer, 'Contracting Around Privacy: The (Behavioral) Law and Economics of Consent and Big Data' (2017) 8(1) JIPITEC 12, observes that for these "algorithms to allow good predictions about personal traits and behaviors, the network operator needs two things: sound knowledge about the social graph [describing the social ties between users] and large amounts of data."

9 Article 29 Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (2017) 17/EN WP 251.

10 See Bart W. Schermer, 'The limits of privacy in automated profiling and data mining' (2011) 27 Computer law & security review 45, 52, and Article 29 Working Party (n 9) 5.

11 Lorraine Daston, 'How Reason Became Rationality' (*Max-Planck-Institut für Wissenschaftsgeschichte*, 2013) <https://www.mpiwg-berlin.mpg.de/en/research/projects/DeptII_Daston_Reason> accessed 1 March 2018.

12 Solon Barocas et al., 'Governing Algorithms: A Provocation Piece' (*SSRN*, 4 April 2013) 9 <https://ssrn.com/abstract=2245322> accessed 1 March 2018.

13 Joshua A. Kroll et al., 'Accountable Algorithms' (2017) 165 U Pa L Rev. 633.

14 Other routes may be explored. In the US, Keats Citron (n 4) 33 suggested that the principles of due process may constitute

(GDPR)[15] bans solely automated decisions having legal effects on the data subject's life "or similarly significantly affects him or her."[16] However, when such decisions are allowed, the data controller shall ensure the transparency of the decision, and give the data subject the rights to obtain human intervention, to express their point of view, as well as to contest the decision. Data protection is the most studied perspective but invoking it by itself is a strategy that "is no longer viable."[17] Therefore, this paper approaches this issue by integrating data protection, intellectual property, and freedom of information.

7 As to the intellectual property route, some copyright and patent exceptions may allow the access to a computer program implementing an algorithm, notwithstanding its proprietary nature.

8 In turn, when it comes to the freedom of information, an Italian court stated that an algorithm is a digital administrative act and therefore, under the freedom of information regime, the citizens have the right to access it.[18]

9 In terms of method, the main focus is a desk-based research of EU laws, and of the UK and Italian implementations. The paper is both positive and normative. Whilst advocating against algorithmic decision-making, this research adopts a pragmatic approach whereby one should take into account that the replacement of human decision-makers with algorithms is already happening. Therefore, it is important to understand how to solve the relevant legal issues using existing laws. If algorithms are becoming "weapons of math destruction,"[19] it is crucial that awareness is raised regarding the pervasivity of algorithmic decision-making and that light is shed on the existing legal tools, in anticipation of better regulations and more responsible modelers. Without clarity on the nature of the phenomenon and the relevant legal tools, it is unlikely that citizens will trust algorithms.

a sufficient answer against algorithmic decisions (in particular, against automated scoring systems). The authors recommend that the Federal Trade Commission interrogate scoring systems under their unfairness authority.

15 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [2016] OJ 119/1.

16 GDPR, art 22.

17 Schermer (n 10) 52.

18 TAR Lazio, chamber III *bis*, 22 March 2017, No 3769.

19 Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Crown 2016).

B. Positive and normative arguments against algorithms as a replacement for human decision-makers

10 The first part of this section is dedicated to presenting the main reasons why algorithms cannot replace human decision-makers. The second part discusses the reasons why such a replacement is not desirable. The analysis is carried out with the judge as the model of a decision-maker.

I. The unfeasibility of the replacement

11 The untenability of the replacement is mainly related to the role and characteristics of legal interpretation. Algorithms could replace human decision-makers if interpretation were a straightforward mechanical operation of textual analysis; where the meaning is easily found by putting together the facts and the norms. The said model of interpretation, which seems flawed, is accompanied by the conviction that there is a clear distinction, on the one hand, between interpretation and application and, on the other hand, between easy cases and hard cases. However, legal interpretation seems to have the opposite characteristics. Indeed, it is ubiquitous[20] and its extreme complexity relates to several factors,[21] such as the psychological (and not merely cognitive) nature of the process.[22] This highlights

20 Given the features of legal interpretation in practice, the brocard *in claris non fit interpretatio* should be replaced by *in claris fit interpretatio* (*cf* Francesco Galgano, *Tutto il rovescio del diritto* (Giuffrè 2007) 100, who points out how the attempts to rule out legal interpretation by means of clear statutes (Carlo Ludovico Muratori) or through proposals to expressly prevent judges from interpreting the statutes (Pietro Verri) today would be laughable. *cf* Vittorio Villa, *Una teoria pragmaticamente orientata dell'interpretazione giuridica* (Giappichelli 2012).

21 For instance, due to the intrinsic vagueness of the legal language and because of the importance of general principles, one of the main tasks of judicial interpretation is striking a balance between conflicting interest, which shall be done on a case-by-case basis. However, some scholars believe that "balancing works with mathematical rules" (Pier Luigi M. Capotuorto, 'Arithmetical Rules for the Judicial Balancing of Conflicts between Constitutional Principles: From the 'Weight Formula' to the Computer-Aided Judicial Decision' (2007) 3(2) Rivista di Diritto, Economia e Gestione delle Nuove Tecnologie 171.

22 Richard A Posner, 'The Role of the Judge in the Twenty-First Century' (2006) 86 B U L Rev 1049, 1060, believes that the psychological component is dominant when it comes to the sources of ideology, which plays a fundamental role in the decisions of all judges. Works on the prediction of judicial decisions usually focus on non-textual elements such as the nature and the gravity of the crime or the preferred policy position of each judge. See e.g. Benjamin E Lauderdale and

why it is currently impossible to develop an algorithm capable of interpreting the law as a human judge would do.[23] The high degree of discretion of the relevant process seems to be the main reason for the impossibility of the replacement. Dworkin's view whereby there is only one right answer to legal questions[24] has very few defenders indeed.[25] Hart[26] clearly proved his doctrine of strong discretion in judicial interpretation, as "a necessary byproduct of the inherent indeterminacy of social guidance."[27] A factor that increases the hermeneutical discretion is that interpreting and applying the law requires value judgements and choices, which are very hard to formalise and compute because of their indeterminacy.[28] One may object that AI may replace humans at least in the legal interpretation of easy cases (for instance, because there is a robust body of case law on the exact issue at hand). However, it has been shown that it is impossible to determine *ex ante* whether a case is easy or difficult: the complexity of the legal experience tells us that the factual and normative circumstances make a case easy or difficult. The similar suggestion to limit the use of algorithms to the application of the law is based, finally, on the wrong assumption that there is an interpretation-application dichotomy and that there is no room for interpretation when one applies the law. Conversely, application seems the last (and most important) phase of the interpretive process.[29]

12 Even leaving the philosophy of law aside, the actual development of statutory interpretation shows the increasing discretion of this activity. Indeed, it seems clear that nowadays the literal rule of interpretation plays a small and often rhetoric role, whereas a purposive approach to statutory interpretation has become commonplace,[30] in part as a consequence of the EU's influence. It has been noted that, whatever the philosophical view one adopts, the discretional power of courts is never expressed in a pure mechanical operation.[31] A good example of the new face of legal interpretation is provided by the case of the Psychoactive Substances Act 2016.[32] The parliamentary debate[33] clearly shows that the intention of the legislator was to ban the so-called poppers (of the class 'alkyl nitrites'), a recreational drug used traditionally by men who have sex with men due to its effects on the relaxation of muscles (including the sphincter). The broad definition of psychoactive substance seemed to allow the interpretation whereby poppers were banned[34] and some law enforcement agencies applied it consistently.[35] However, the final result is that

Tom S Clark, 'The Supreme Court's many median justices' (2012) 106(4) American Political Science Review 847.

23 There are several studies in the field of AI & Law that develop models to explain the legal reasoning, but this is an ex-post operation, as opposed to a genuinely predictive one. See, for instance, Latifa Al-Abdulkarim et al., 'A methodology for designing systems to reason with legal cases using Abstract Dialectical Frameworks' (2016) 24(1) Artif Intell Law 1.

24 See, for instance, Ronald Dworkin, 'No Right Answer?' (1978) 53 New York University Law Review 1; Ronald Dworkin, *Law, Morality, and Society* (Peter Hacker & Joseph Raz eds, Clarendon Press 1977).

25 For instance, an author like Thomas B. Colby, a strong assertor of the rule of law, recognises that the law is often ambiguous or open-ended and, therefore, "there is no objectively correct answer that can be discerned simply by calling balls and strikes." (Thomas B. Colby, 'In Defense of Judicial Empathy' (2012) 96 Minn. L. Rev. 1944, 2015) Even those who argue for an overcoming of the centrality of the Hart-Dworkin debate cannot "envision a jurisprudential future without Hart's masterful work at its center" (Brian Leiter, 'Beyond the Hart-Dworkin Debate: The methodology problem in jurisprudence' (2003) 48 Am. J. Juris. 17, 18). For a recent critique to Dworkin, see Aulis Aarnio, 'One right answer?' [2011] Essays on the doctrinal study of law 165. As suggested by Tony Ward in his comments on a previous draft of this paper, one should note that, even in the event that Dworkin were right, it is unclear how the Hercules algorithm would be programmed.

26 See H. L. A Hart, *The Concept of Law* (Penelope Bulloch & Joseph Raz eds, Clarendon Press 1994) 123.

27 Scott J. Shapiro, 'The "Hart-Dworkin" debate: A short guide for the perplexed' (2007) Public Law and Legal Theory Working Paper Series No. 77, 16.

28 It has been noted that "even if computers were technically able to mimic legal decision making in a mechanical fashion they would necessarily miss the subtle institutional, value-based, experiential, justice-oriented, and public policy dimensions that are the heart of lawyerly analysis" (Lisa A. Shay et al., 'Do robots dream of electric law? An experiment in the law as algorithm' in Ryan Calo, A. Michael Froomkin, and Ian Kerr (eds) *Robot Law* (Elgar 2016) 274, 277, citing Harry Surden, 'Computable Contracts' (2012) 46 U.C. Davis L. Rev. 629).

29 Francesco Viola, 'Interpretazione e indeterminatezza della regola giuridica' (2002) 7-8 Diritto privato 49, 51, explains in this way one of the differences between Hart and Kelsen (the *ex-ante* distinction between easy and difficult cases – and between interpretation and application – would be possible adopting a Kelsenian perspective).

30 Catherine Elliott & Frances Quinn, *English Legal System* (17th ed, Pearson 2016) 61. The shift is very significant, and it is suggested already from the use of 'approach' instead of 'rule', which hints at a more flexible strategy drawing from several sources and taking into account many factors, rather the mechanical operation of subsuming a fact under a rule.

31 Pier Luigi M. Lucatuorto, 'Modelli computazionali della discrezionalita del giudice: uno studio preliminare' (2006) 7(3) Ciberspazio e diritto 1, 2.

32 I am thankful to Chris Ashford for the insight provided in his "The UK Poppers 'Ban' and the Psychoactive Substances Act 2016: New Legal Frontiers in the Homonormative Imagination" (Northumbria University Gender, Sexuality and Law Research Seminar, Newcastle upon Tyne, 14 June 2017).

33 The Burnham amendment and the Scottish National Party (SNP) one, aimed to allow poppers, were rejected.

34 Psychoactive Substances Act 2016, s 2.

35 See Steven Hopkins, 'Crawley Police Forced To Apologise After Wrongly Seising Poppers After Legal High Ban Came Into Effect' (*Huffington Post UK*, 26 May 2016) <http://www.huffingtonpost.co.uk/entry/crawley-police-forced-to-

poppers *are not* banned, because the UK Advisory Council on the Misuse of Drugs explained that since poppers have a merely indirect effect on the nervous system, they do not technically qualify as psychoactive substances and, therefore, fall outside the scope of the Act.[36] Finally, sectoral empirical studies[37] are showing that algorithms cannot cope with legal interpretation in a satisfactory way. For instance, it has been shown[38] that algorithms often reflect a wrong interpretation of the law they enforce,[39] in particular with regards to the fair use analysis in online infringement cases.[40] These are just a couple of examples of how interpretation is discretionary, ubiquitous, complex, and unpredictable.[41] Therefore, it seems that it is currently impossible to design an interpretive algorithm.

13 This study itself confirms this view, in as much as from an apparently simple provision, such as Article 22 of the GDPR, stem a number of complicated interpretative problems for which there is no easy answer. The relevant difficulties will be explained in section 4 below. Here suffice to say that there is a meta-problem. Even if algorithms could perfectly replace human decision-makers, arguably it would not be fair to let them interpret a provision – Article 22 – which has the aim of protecting citizens from algorithmic decisions.

14 The above considerations regard the current progress in algorithms-related technologies. However, AI's growth is exponential, therefore the considerations above may prove to be wrong soon, especially in fields where the issues arising are often similar and there is a lot of precedent. Less so where there is no established case law, and/or the field is fast evolving.[42] For example, predicting the outcome of succession cases involving only land may prove easier than cyber law cases with cross-border elements. That said, alongside the technologies, the scholarship is evolving. Recently, the first systematic study on predicting the outcome of cases tried by the European Court of Human Rights based solely on textual content was presented.[43] The model is quite accurate, being able to predict the outcome in 79% of cases. However, there are some considerations to be made. Especially in matters as important as human rights, reaching a wrong decision in 21% of the cases would be utterly unacceptable. Secondly, the reasons for this margin of error should be better analysed; they might stem from the fact that interpretation is not a mere mechanical operation of text analysis. Thirdly, the authors themselves point out that the model would not be a substitute for the human decision-maker, because its role would rather be an "assisting tool"[44] to identify cases and extract patterns. Lastly, the study did not predict the outcome using the documents filed by the applicants, but only analysing the published rulings. This means that a human judge had already selected the materials and interpreted them, which affects the results of the study.[45] More generally, it still holds true that "[j]ustification, persuasion and discretion are the main limits of the Artificial Intelligence application in Law."[46]

15 Second, human learning is much more complex than machine learning. According to the seminal *Mind over Machine*,[47] the characteristics of human learning would explain why prophecies about real machine intelligence have all been proven wrong,[48] and why

apologise-after-wrongly-seising-poppers-after-legal-high-ban-came-into-effect_uk_57472a41e4b0ebf6a3297cac> accessed 1 March 2018.

36 Advisory Council on the Misuse of Drugs, 'CMD review of alkyl nitrites (poppers)' (*The UK Government,* 16 March 2016) <https://www.gov.uk/government/publications/acmd-review-of-alkyl-nitrites-poppers> accessed 1 March 2018.

37 Along with the studies cited in the following footnotes, see, for instance, Joe Karaganis & Jennifer Urban, 'The rise of the robo notice' (2015) 58(9) Communications of the ACM 28.

38 Maayan Perel & Niva Elkin-Koren, 'Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement' (2017) 69 Fla. L. Rev. 181, 210.

39 Kenneth A. Bamberger, 'Technologies of Compliance: Risk and Regulation in a Digital Age' (2010) 88 Tex. L. Rev. 669, 675-6.

40 Specifically, when the researchers tried to upload a 48 seconds homemade video of a child dancing a protected song by Justin Bieber, 25% of video-sharing platforms removed the video, notwithstanding the fact that it clearly constituted a fair use.

41 Reed C. Lawlor, 'What Computers Can Do: Analysis and Prediction of Judicial Decisions' (1963) 49(4) American Bar Association Journal 337, conjectured that in the future machines would be able to predict the outcomes of judicial decisions. The prophecy was not fulfilled (yet), but Nikolaos Aletras et al., 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective' (2016) 2:e93 PeerJ Comput. Sci., constitute a progress.

42 However, as said above, it is not possible to assess *ex ante facto* whether a case is easy or hard (and even *ex post facto*, the lines are blurred and interpretation is needed for the hard cases as well as for the easy cases).

43 Aletras (n 41).

44 ibid 3.

45 ibid 2, assume that "the text extracted from published judgments of the Court bears a sufficient number of similarities with, and can therefore stand as a (crude) proxy for, applications lodged with the Court as well as for briefs submitted by parties in pending cases". They accept, however, that "full acceptance of that reasonable assumption necessitates more empirical corroboration".

46 Pier Luigi M. Lucatuorto, 'Computer-Aided Sentencing: Computer Science and Legal Aspects: The Chinese Case' (2006) 2(4) Rivista di Diritto, Economia e Gestione delle Nuove Tecnologie 388.

47 Hubert L Dreyfus and Stuart E Dreyfus, *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer* (Free Press 1988).

48 For instance, Herbert A Simon, *The Shape of Automation for*

small scale successful experiments conducted in laboratories were not as successful once extended to larger systems and the real world. In particular, machines will not be able to replace human beings when cognitive tasks require intuition and holistic thinking.[49] By presenting a five-stage model of acquisition of expertise (novice, advanced beginner, competent, proficient, and expert), these authors show that there is more to human intelligence than the computer's calculative rationality. Only the human brain, at least currently, is capable to properly learn and understand through holistic intuition a world that is – unlike the laboratory – incomplete, imprecise, and unreliable. It seems, indeed, unlikely that training a machine with millions of legal provisions and case law can lead to the same results to the learning of a judge, who is immersed in the real world and learns in ways, which cannot be coded.

II. Eight arguments against the desirability of algorithms replacing human decision-makers

16 Let us assume that the thesis of this paper is wrong. Let us say, for the sake of argument, that either interpretation is not ubiquitous, or it is not discretionary (or that algorithms can cope well with strongly discretionary processes). Let us posit, then, that algorithms can learn in the same way as the humans. Nonetheless, there are at least eight reasons why they *should not* replace human decision-makers. Two reasons refer to why one should not trust algorithms. Six arguments are, in turn, presented to show why we should trust humans.

1. The replacement is undesirable because there are good reasons not to trust the algorithms

17 Let us start with what is not to like in algorithms. One of the strong arguments in favour of the algorithms is that they are more reliable than human beings are. However, there is evidence that algorithms can make mistakes and, when they do so, the effects are on a larger scale than an error made by a human judge in a ruling.[50] More importantly, algorithms are not more reliable than human beings, because of the emerging problem of algorithmic (or machine) bias.[51] The founder of the Algorithmic Justice League, for instance, stated that a facial recognition machine could not see her because she is black and, probably, the machine learning algorithm was trained only using white faces.[52] Contrary to popular belief, algorithms do not eliminate bias, because the relevant models are opaque, unregulated, and incontestable.[53] Even those who believe that AI should be used (in combination with law and self-regulation) for the governance of the Internet, admit that the "[l]ack of transparency on how algorithms operate is a real issue, as well as the problem that artificial intelligence tends to share the biases of the humans it learns from."[54]

18 In the context of the UK inquiry on algorithms in decision-making,[55] six reasons why algorithmic systems can produce biased outcomes have been presented.[56] First, design choices make the decision-making process or the factors it considers too opaque; these choices may also limit the control of the designer.[57] Second, the output of the system may

Men and Management (Harper & Row1965) 38, foresaw that in 1985 machines would have been capable of doing any work that a man could do. In hindsight, that prediction was not entirely accurate.

49 Computer "reasoning" is deemed to be ontologically different to human know-how: "a far superior holistic, intuitive way of approaching problems that cannot be imitated by rule-following computers" (Dreyfus (n 47) 193). For some recent developments in intuition modelling, see Ulrich Hoffrage and Julian N Marewski, 'Unveiling the Lady in Black: Modeling and aiding intuition' (2015) 4 Journal of Applied Research in Memory and Cognition 145.

50 One need only think of the wrong calculations that affected 20.000 divorced couples due to a software glitch (see, e.g. Will Grice, 'Divorce error on form caused by UK Government software glitch could affect 20,000 people' (*The Independent*, 18 December 2015) <http://www.independent.co.uk/news/uk/home-news/ministry-of-justice-software-glitch-could-see-thousands-revisiting-painful-divorce-settlements-a6777851.html> accessed 1 March 2018.

51 Cf. Megan Garcia, 'Racist in the Machine: The Disturbing Implications of Algorithmic Bias' (2016) 33(4) World Policy Journal 111, and, more generally, Kroll (n 13) 633.

52 <http://www.ajlunited.org/> accessed 1 March 2018. These kind of problems had already been evidenced by Brandan F. Klare et al., 'Face Recognition Performance: Role of Demographic Information' (2012) 7(6) IEEE Transactions on Information Forensics and Security 1789.

53 This is one of main ideas of O'Neil (n 19).

54 Andrés Guadamuz, 'Whatever happened to our dream of an empowering Internet (and how to get it back)' (*TechnoLlama*, 5 June 2017), <http://www.technollama.co.uk/whatever-happened-to-our-dream-of-an-empowering-internet-and-how-to-get-it-back> accessed 1 March 2018. On the phenomenon of machine bias (or algorithmic bias) see below.

55 <https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/inquiries/parliament-2015/inquiry9/> accessed 1 March 2018.

56 Science and Technology Committee, 'Written evidence submitted by Dr Alison Powell (ALG0067)' (*UK Parliament*, 2017) <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/algorithms-in-decisionmaking/written/69121.html> accessed 1 March 2018.

57 Alan Dix, 'Human issues in the use of pattern recognition techniques', in R Beale and J Finlay (eds), *Neural Networks*

be affected by the biases in data collection.[58] Third, unlike human beings, algorithms cannot balance biases in interpretation of data by a conscious attention to the redress of the bias.[59] Fourth, there are biases in the ways that learning algorithms are tuned based on the testing users' behaviour.[60] Fifth, algorithms may be designed for a purpose, but then inserted into systems designed for other purposes.[61] Lastly, as already said with regard to the Algorithmic Justice League, another factor is the biases in the data used to train the decision-making systems.[62]

19 Algorithmic bias is the main problem regarding automated decision-making with legal effects.[63] It has been submitted that "while persistent inequities stem from a complex set of factors, digitally automated systems may be adding to these problems in new ways."[64] It is arguable that even if the automated decision (e.g. a ruling) is biased, the move to algorithms "may at least have the salutary effect of making bias more evident."[65] Algorithmic bias is dealt with in a recital of the GDPR,[66] in a way which is not entirely satisfactory. Indeed, the GDPR calls on the data controller to "use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimised, secure personal data in a manner that takes account of the potential risks involved for the interests and rights of the data subject and that prevents, inter alia, discriminatory effects"[67] on the basis of sensitive data. Now, it would seem that the GDPR's focus is misplaced. The point with discrimination is not only that the data are inaccurate or that they are not secure. The main problem is that these data should never be used to discriminate in the first place,[68] regardless of their being accurate or not, or that independency constraints should be put in place.[69]

20 The second argument revolves around transparency. Indeed, making bias evident would mean ensuring transparency, which seems a chimera for a number of reasons, including the fact that the more accurate an algorithm is, the less transparent.[70] The trade-off accuracy vs. transparency is easily explained. On the one hand, modelers tend to develop more accurate models "with increasingly complex, data-mining-based black-box models."[71] On the other hand, model users tend to favour "transparent, interpretable models not only for predictive decision-making but also for after-the-fact auditing and forensic purposes."[72] Against the dominant idea that transparency will solve all the problems, some scholars point out that "[d]isclosure of source code is often neither necessary (because of alternative techniques from computer science) nor sufficient (because of the issues analysing code) to demonstrate the fairness of a process."[73] Arguably, however, such disclosure would be necessary to comply with the right to an effective remedy and to a fair trial under the EU Charter of Fundamental Rights and the European Convention on Human Rights.

21 The lack of transparency is related to the so-called black box (better said, black *boxes*). Arguably, three different black boxes may be distinguished: the organisational; the technical; and the legal one. The organisational black box will not be the subject of specific analysis. Suffice to say that algorithms are mostly implemented by "private, profit-maximising entities, operating under minimal

and Pattern Recognition in Human Computer Interaction (Ellis Horwood 1992) 429.

58 Stella Lowry & Gordon Macpherson, 'A blot on the profession' (1988) 296 British Medical Journal 657.

59 Aylin Caliskan et al., 'Semantics derived automatically from language corpora contain human-like biases' (2017) 356(6334) Science 183.

60 Dix (n 57) 57.

61 Louise Amoore, *The politics of possibility: risk and security beyond probability* (Duke University Press 2013).

62 Klare (n 52).

63 Algorithmic bias has many potential consequences. For instance, in the context of social media, it may lead to the so-called filter bubble. See, e.g., William H. Dutton et al., 'Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States' (*Quello Center Working Paper No. 5-1-17*, 2017).

64 Seeta Peña Gangadharan et al., *Data and Discrimination: Collected Essays* (Open Technology Institute 2014).

65 Barocas (n 12) 9.

66 GDPR, recital 71.

67 ibid

68 Rakesh Agrawal and Ramakrishnan Srikant, 'Privacy-Preserving Data Mining' (2000) 29(2) ACM SIGMOD Record 439 (2000).

69 Toon Calders et al., 'Building Classifiers with Independency Constraints' (2009) IEEE ICDM Workshop on Domain Driven Data Mining 13. Therefore, for instance, sensitive attributes such as sex shall be included, but the program would be instructed to make predictions independently of the said attributes. This second approach seems preferable for accountability reasons.

70 Barocas (n 12) 9 accept that "algorithms may involve rules of such complexity that they defy attempts to trace their reasoning".

71 Innocent Kamwa et al., 'On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors' (2012) 3(1) IEEE Transactions on smart grid 152.

72 ibid 152. They conclude that "for catastrophe anticipation purposes, we would favor fuzzy logic-based transparent solutions over black box solutions for implementation ease and robustness, as well as for their suitability in the auditing process, even while sacrificing some predictive accuracy" (ibid 160).

73 Kroll (n 13) 633.

transparency obligations."[74] As to the technical black box, artificial intelligence makes the rationale of decisions intrinsically difficult to access. This is particularly evident with the so-called neural networks that, being modelled on the brain, are at least as opaque. One need only imagine a deep-learning neural network which is trained using old mammograms that have been labelled according to which women went on to develop breast cancer.[75] It could help us to make predictions on which breasts are likely to develop cancer, but without knowing the risk factors (the rationale), it is unlikely that the patient would undergo therapy and, more generally, the development of cancer research would not be substantive. The legal black box relates to intellectual property and will be presented in the following section.

22 The lack of transparency has obvious repercussions on the accountability issue. For instance, ensuring fair, lawful, and transparent processing may be difficult "due to the way in which machine learning works and / or the way machine learning is integrated into a broader workflow that might involve the use of data of different origins and reliability, specific interventions by human operators, and the deployment of machine learning products and services."[76] Some technical tools to ensure accountability in algorithmic scenarios have been presented,[77] but they do not seem sufficient to offset the inherent problems in algorithmic decision-making.

2. The replacement is undesirable because there are good reasons to trust the human beings

23 This subsection is dedicated to the reasons why one should trust humans over algorithms and, more generally, over non-human agents.

24 First, human beings tend to emulate the behaviour of the majority of fellow human beings. This should ensure consistency and predictability in societal behaviours. This phenomenon was observed with particular clarity by Solomon Asch, who developed the so-called psychology of conformity.[78] Needless to say that non-human agents do not have a consciousness[79] and, therefore, psychology does not apply to them. One could object, however, that conforming to the majority does not equal pursuing the common good, because it could lead to the oppression of the minorities. However, humans have some built-in safeguards.

25 The argument can be put forward that, despite the different characteristics of human beings, humans tend to act consistently towards the common good. This may be explained with the power of sanctions.[80] Human beings comply with the law not for a natural disposition, but because they do not wish to be sanctioned. However, it is hardly arguable that non-human agents share this fear. Indeed, neither can they be imprisoned (criminal sanctions), nor do they own assets that can be used to execute civil and administrative sanctions.

26 The third argument refers, like the previous one, to the effects of group pressure, but in a different setting. It can be summed up by saying that hypocrisy has a civilising force.[81] Indeed, with regards to the relationship between deliberation and publicity, it has been observed that "the effect of an audience is to replace the language of interest by the language of reason and to replace impartial motives by passionate ones."[82] These considerations, rooted in human psychology, do not apply to non-human agents. Therefore, hypocrisy cannot civilise algorithmic decision-makers.

27 Let us say that it is possible for an algorithm to learn and decide like a human judge. At this point, one

74 Perel & Elkin-Koren (n 38) 181.

75 The scenario, imagined by Andrea Vedaldi (University of Oxford) is referred to by Davide Castelvecchi, 'Can we open the black box of AI?' (*Nature*, 15 October 2016) <http://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731> accessed 1 March 2018. *cf* Krzysztof J Geras, Stacey Wolfson, Yiqiu Shen, S. Gene Kim, Linda Moy, and Kyunghyun Cho, 'High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks' (2017) arXiv:1703.07047 [cs.CV].

76 Dimitra Kamarinou et al., 'Machine Learning with Personal Data' (*Queen Mary School of Law Legal Studies Research Paper No. 247*, 2016) 22. Christopher Kuner et al., 'Machine learning with personal data: is data protection law smart enough to meet the challenge?' (2017) 7(1) International Data Privacy Law 1, observe that "[m]achine learning is data driven, typically involving both existing data sets and live data streams in complex training and deployment workflow [therefore it] may be difficult to reconcile such dynamic processes with purposes that are specified narrowly in advance."

77 Kroll (n 13) 633.

78 See Solomon E Asch, 'Effects of group pressure upon the modification and distortion of judgment', in H Guetzkow (ed), *Groups, leadership and men* (Carnegie Press 1951) 222.

79 However, there is a significant debate about artificial consciousness, whose functions have been described as including awareness of self, will, instinct, and emotion (Igor Aleksander, 'Machine consciousness' (2008) 3(2) Scholarpedia 4162). It seems that the prevalent position in the literature is against the existence of a proper artificial consciousness.

80 *cf* Antonio Pagliaro, 'Sanzione. Sanzione penale' [1992] 28 Enciclopedia giuridica 3; David R Carp, 'The Judicial and Judicious Use of Shame Penalties' (1998) 44(2) Crime & Delinquency 277.

81 Jon Elster (ed), *Deliberative Democracy* (Cambridge University Press 1998).

82 ibid 111.

may argue, it would be sufficient to find the best judge in the world and create a large number of non-human clones that will gradually replace all human judges. However, this scenario raises some issues. Pluralism seems to be the main one.[83] Indeed, if pluralism is rooted in the respect for the minorities and in the belief that a multiplicity of viewpoints enriches the understanding of the world, then erasing this by cloning the perfect judge would at least be problematic. Even before that, how does one find the perfect judge to clone? What does it mean to be the best judge? Is it possible to entirely eliminate human bias?[84]

28 A fifth reason why this paper takes a humanist stance is empathy, which is the "cognitive ability to understand a situation from the perspective of other people, combined with the emotional capacity to comprehend and feel those people's emotions in that situation."[85] This could come as a surprise, since usually empathy is seen as a bias[86] and, therefore, as an argument in favour of non-human agents. Conversely, empathy is "a requirement of judicial neutrality."[87] It has been shown that arguments in favour of judicial empathy are rooted, perhaps unexpectedly, in "a firm commitment to the rule of law and a deep-seated appreciation of—rather than rejection of— legal doctrine."[88] A recent study shed light on the shortcomings of the anti-empathic consensus; indeed, it descends of XIX century formalism, but it has "drifted from its source such that it would almost certainly be condemned by the very formalist scholars from whom it is descended."[89] Not only is empathy not a defect in human decision-making, it serves a positive function. This is required by the paramount function of concepts such as reasonableness and balancing tests.[90] More generally, it can be argued that empathy is the way justice (as opposed to law) enters the decision. When Cicero wrote *"summum ius summa iniuria"*[91] he meant that the mechanical application of the law leads to unjust results. Empathy tempers legalistic excesses and algorithms are not capable of it.

29 Lastly, one needs to choose between democracy and technocracy. In a democratic context, laws are the product of a debate between politicians. This debate is public, and the politicians are democratically elected and accountable both politically and legally. Human judges are either democratically elected or receive specific legal training. Conversely, algorithmic law (as in Lessig's "code is law"[92]) is more problematic. Indeed, "software development, even open source, is opaque, and concentrated in a small programming community, many of whom are employed by few oligopolistic corporations directly accountable to no external party."[93] Algorithms could be suitable to apply algorithmic laws, but given the said characteristics, it is hoped that their role and scope remains limited.

30 For the reasons above, the replacement of algorithms to human beings seems both unfeasible and undesirable.

C. Intellectual property rights: more a problem, than a solution

31 Even though there are good reasons to believe that algorithms cannot and should not replace human decision-makers, it is becoming obvious that the replacement is already taking place, regardless of the relevant pitfalls. Therefore, a lawyer should be able to provide a sufficiently clear answer to a client subject to an algorithmic decision.

32 There are at least three routes that can be taken, should the relevant requirements be met. In this section, the focus will be on intellectual property and the relevant exceptions that may enable access to a computer program implementing an algorithm, or the relevant invention, notwithstanding its proprietary nature. The features of the analysed exceptions made scholars talk of "the advent of a more active approach to copyright exceptions,"[94] which creates quasi-rights, "legal hybrids between exceptions and rights."[95] This must be taken into account when interpreting the relevant provisions

83 When asked about this argument during the conferences cited in the acknowledgments, the audience also mentioned other negative repercussions. The most relevant one seems to be the lack of legal innovation deriving from a single approach to decision-making.

84 As to the last question, it is submitted that if one eliminates ideologies in the attempt of eliminating bias, the output would be a useless algorithm, incapable of deciding. Indeed, ideologies guide human judges in deciding, for instance, whether intellectual property rights should prevail over access to knowledge, whether the reasons of privacy should take precedence over those of free speech, etc.

85 Colby (n 25) 1945.

86 See, e.g., Adam N Glynn and Maya Sen, 'Identifying Judicial Empathy: Does Having Daughters Cause Judges to Rule for Women's Issues?' (2015) 59(1) American Journal of Political Science 37.

87 Colby (n 25) 2015.

88 ibid 1946.

89 Brenner Fissel, 'Modern critiques of judicial empathy: A revised intellectual history' (2016) Mich. St. L. Rev 817.

90 Colby (n 25) 1946.

91 Cicero, *De officiis*, I, 10, 33.

92 Lawrence Lessig, *Code* (v.2.0, Basic Books 2006).

93 Kieron O'Hara, 'Smart contracts – Dumb idea' (2017) The Digital Citizen 2, 5.

94 Tatiana-Eleni Synodinou, 'The lawful user and a balancing of interests in European copyright law' (2010) 41 IIC 819, 826.

95 ibid 826.

and striking a balance with the restricted acts. Equally, defences to patent infringement will be dealt with, although there is not enough evidence to claim their nature as quasi-rights.

33 A major issue is understanding the rationale of algorithmic decisions. This is made difficult by the so-called black boxes. The organisational black box and the technical one have been presented above. The legal black box remains to be analysed. This depends primarily on the (ab)use of intellectual property rights (trade secrets, database rights, etc.) and the kindred rights that companies are collecting on the users' data, that do not fit easily in the traditional intellectual property categories and are leading to the datafication of the digital economy. Along the same lines, it has been noted that "data, originating from users, from devices, sent through the 4G and 5G networks to the client servers and the Cloud are heavily boxed in by intellectual property rights."[96]

34 Even though there are many open-source machine learning frameworks (e.g. Apache Singa, Shogun, and TensorFlow), most AI algorithms are proprietary (Google search and Facebook news feed are the classical examples) i.e. covered primarily by trade secrets,[97] which is the "most common form of protection used by business."[98] Under the new Trade Secrets Directive,[99] algorithms can be covered by trade secrets because they are not generally known or easily accessible and they have commercial value.[100] This is true as long as the person who has control of the algorithm takes steps to keep it secret.[101] The general rule is that the unauthorised acquisition, use, or disclosure of algorithms covered by trade secrets is unlawful.[102] However, the acquisition shall be lawful in a limited number of circumstances, the most relevant of which seems to be the "observation, study, disassembly or testing of a product or object that has been made available to the public or that is lawfully in the possession of the acquirer of the information who is free from any legally valid duty to limit the acquisition of the trade secret."[103] This appears to be a reference to one of the permitted uses of computer programs under the Software Directive.[104] There is a potential contrast between the two regimes. To say that the acquisition is legal only if "free from any legally valid duty to limit [it],"[105] may be construed as meaning that if the owner of the algorithm contractually restricts the said exception, then no observation, study, disassembling, or testing of the algorithm would be allowed. However, under the Software Directive, there is a right to "observe, study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program."[106] This Directive goes on pointing out that any contractual provisions contrary to said exception "shall be null and void."[107] In the UK, the Copyright, Designs and Patents Act 1988 is clear where it provides that "it is irrelevant whether or not there exists any term or condition in an agreement which purports to prohibit or restrict the act (such terms being [...] void)."[108] The leading case on the matter is *SAS Institute v World Programming*, where it was found that copyright owners cannot restrict the purposes for which the analysed permitted acts are carried out. Additionally, even though only lawful users can avail themselves of the defence, these are not limited to those who click through the licence.[109]

96 Bjorn Lundqvist, 'Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet of Things World' (Faculty of Law, University of Stockholm Research Paper No. 1/2016) 10.

97 This has the potential to impact many fundamental rights, such as the one of access to public information. For instance, crashes such as the one that, on 6 May 2010, caused the Dow Jones Industrial Average to drop by 9% thus burning millions of dollars, cannot be explained "not least because many of the algorithms involved are proprietary" (Scholz (n 3) 103).

98 James Pooley, 'Trade secrets: The other IP right' (2013) 3 WIPO Magazine.

99 Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure (Trade Secrets Directive). Member States shall transpose it by 9 June 2018. By that date, the UK will still be part of the EU, but prima facie there are not significant differences between the rules on the breach of confidentiality and the new EU regime.

100 Trade Secrets Directive, art 2(1)(a)-(b).

101 Trade Secrets Directive, art 2(1)(c).

102 Trade Secrets Directive, art 4.

103 Trade Secrets Directive, art 3(1)(b).

104 Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Software Directive).

105 Trade Secrets Directive, art 3(1)(b).

106 Software Directive, art 5(3).

107 Software Directive, art 8(2). See also recital 16.

108 Copyright, Designs and Patents Act 1988, ss 50BA and 296A. The Copyright (Computer Programs) Regulations 1992, SI 1992/3233 inserted s 50B in the Copyrights, Designs and Patents Act 1998, allowing lawful users of computer programs to decompile programs to achieve interoperability. In turn, the Copyright and Related Rights Regulations 2003 introduced s 50BA, regarding the exception on observing, studying and testing computer programs.

109 *SAS Institute Inc v World Programming Ltd III* [2013] EWHC 69 (Ch), [60]-[61]. In that case, employees who did not click through the licence had observed and studied the computer programme without infringing copyright because the colleague who acquired the programme was operating on behalf of the employer, which was a legal person. It was deemed immaterial that the licence openly restricted the use to the person who clicked through the licence. At a closer look, the distinction between licensed employee and unlicensed ones. Indeed, art 9(1) of the Software Directive renders null and void any contractual restrictions to the exceptions and "this includes a contractual restriction on the employees by whom a legal person in the position of

35 Moreover, the Trade Secrets Directive itself recognises the legality of the acquisition, use or disclosure of trade secrets for purposes of freedom of expression and information.[110] Arguably, there is not an actual conflict here. As an example, let us imagine one buys an Amazon Echo. Under one of the several contracts that one has to accept, one agrees that "all Confidential Information will remain [Amazon's] exclusive property"[111] and one may not "reverse engineer, decompile, or disassemble"[112] the Alexa[113] Service or the Alexa Materials.[114] Under the Trade Secrets Directive this section would be enforceable; however, since the Software Directive, being a *lex specialis,* will prevail the section would be unenforceable.[115] Indeed, the conflict is merely ostensible.

36 In the event that trade secrets were deemed to prevail over the exceptions provided by the Software Directive, it may be worth it to take account of the relevant defences. The most relevant and flexible defence seems the public interest one. It has been stated that "the right of confidentiality, whether or not founded in contract, is not absolute. That right must give way where it is in the public interest that the confidential information shall be made public."[116] It is noteworthy that the disclosure may be seen as in the public interest if there has been non-compliance with a legal obligation.[117] One may argue that the circumvention of the Software Directive consisting in secreting an algorithm in an absolute way falls within this scenario. However, the defendant in the relevant infringement proceedings would need to prove that the disclosure be in the public interest and not merely interesting to the public, which may be difficult.[118] Unfortunately, the Trade Secrets Directive does not leave much room for the public interest or other defences. However, it recognises that the Directive shall not affect "the application of [EU] or national rules requiring trade secret holders to disclose, for reasons of public interest, information, including trade secrets, to the public or to administrative or judicial authorities for the performance of the duties of those authorities."[119] The European provision regarding the exceptions does not introduce a stand-alone public interest defence. Indeed, a defence is available if the acquisition, use, or disclosure was "for revealing misconduct, wrongdoing or illegal activity, provided that the respondent acted for the purpose of protecting the general public interest".[120] Unlike the EU, in the UK the public interest is a defence in its own right.[121] Since the transposition deadline is in June 2018, one needs to wait and see how this provision will be interpreted.[122]

37 Additionally, one should remember that copyright protection covers both source code and object code[123] of the computer program implementing the algorithm. However, it leaves out some aspects, such as functionalities, data file formats, programming language, and graphic user interface. They are treated as "ideas" and therefore not copyrightable due to the idea-expression dichotomy.[124] The dichotomy is

WPL can exercise the right under Article 5(3)" (ibid [61]). The decision was upheld in appeal, although for different reasons (*SAS Institute v Worlds Programming Ltd IV* [2013] EWCA Civ 1482 [61], [109] per Tomlinson LJ).

110 Trade Secrets Directive, art 5(a). However, this defence risks weaknesses because the courts may be tempted to interpret it narrowly given that the underlying debate was about the protection of whistle-blowers and journalists, as one can also infer from the express reference to the freedom and pluralism of media.

111 Alexa Voice Service Agreement (last updated 30 January 2017), s 8. The Agreement was updated on 15 February 2018 in order to add Amazon Seller Services Pvt Ltd as the Amazon Party to the Agreement for developers who reside in India. Since the update is minor, this paper keeps referring to the previous version.

112 Alexa Voice Service Agreement, s 9.

113 Amazon's AI virtual assistant.

114 These include "images, audio, logos, specifications, code, documents, data, software, software development kits, libraries, application programming interfaces, applications, services and other information, technology, and related materials" (Alexa Voice Service Agreement, s 2).

115 In *SAS Institute Inc v Worlds Programming Ltd I (WPL)* [2010] EWHC 1829 (Ch), a similar program had been developed studying the competitor's one in breach of the license terms, because the purpose of the permitted act was not learning to use the SAS system (the sole purpose allowed by the license). After the reference to the Court of Justice and *SAS Institute Inc v World Programming Ltd II* [2012] ECR, the national court stated that if an act (e.g. studying) is permitted by the license, the purpose thereof is immaterial, and the exception operates (*SAS Institute III* (n 109); *SAS Institute IV* (n 109) [101]. On the EU case, see Guido Noto La Diega, 'Le idee e il muro del suono: I programmi per elaboratore nella più recente giurisprudenza europea' (2013) 2 Europa e diritto private 543.

116 *Campbell v Frisbee* [2002] EMLR 31, [23].

117 Lionel Bently and Brad Sherman, *Intellectual Property Law* (4th ed, OUP 2014) 1181.

118 *Lion Laboratories Ltd v Evans* [1985] QB 526, 537.

119 Trade Secrets Directive, art 1(2)(b).

120 Trade Secrets Directive, art 5(b).

121 *Initial Services v Putterill* [1968] 1 QB 396.

122 The only Member State that transposed the Trade Secrets Directive is Croatia, with *Zakon o zaštiti neobjavljenih informacija s tržišnom vrijednosti* ('Law on the Protection of Unpublished Information with Market Value') of 30 March 2018. Regrettably, not only the Croatian statute does not introduce a stand-alone public interest defence: it introduces an exception which is narrower than the Directive, being seemingly reduced to a defence for journalists (see art 8(1) and its reference to reporting, media, and pluralism).

123 Agreement on the trade-related aspects of intellectual property rights (TRIPs), art 10(1).

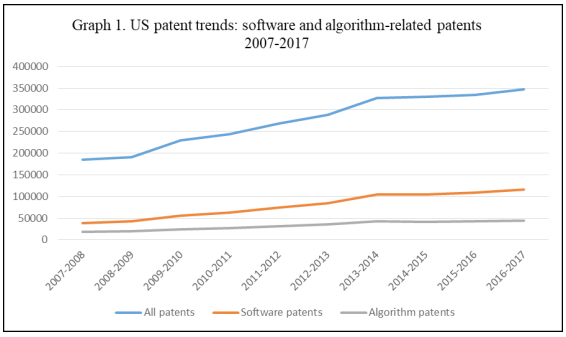
124 In the field of computer programs, on the idea whereby copyright covers the expression of the ideas and not the ideas in themselves, see *SAS Institute II* (n 117); *Navitaire Inc*

also one of the alleged reasons of the patentability of computer-implemented inventions. It has been noted, indeed, "copyright is not a sufficient form of protection where it is the *idea* behind the program which is its commercially valuable element."[125] Computer-related inventions are growing significantly also in connection to the Internet of Things,[126] despite the fact that the relevant patents can stifle innovation.[127] In the US,[128] in September 2017, there were 481,608 patent specifications referring to algorithms.[129] More than 67% of the algorithm-related patents (325,805) were issued over the last ten years with a growing trend reflecting the general increase in patents as shown by Table 1 and Graph 1.[130] Nearly 13% of all patents granted over the last 12 months concern algorithms (ten years ago only 9% of patents were algorithm-related).

38 Table 1. Software and algorithm patent trends in the US (2007-2017).

Period	All patents granted	Software patents	Algorithm patents
2016-2017*	346,543	115,896	44,110
2015-2016	333,767	108,305	42,481
2014-2015	329,722	104,212	41,125
2013-2014	327,729	103,918	42,215
2012-2013	288,989	84,891	35,427
2011-2012	268,157	74,689	32,070
2010-2011	244,338	63,638	27,377
2009-2010	229,694	56,367	24,920
2008-2009	190,285	42,947	19,426
2007-2008	185,340	38,225	17,871

** The period analysed is from 9 September 2016 to 9 September 2017. The same applies to the following rows.*



39 In theory, in the countries that signed the European Patent Convention (and in the others which adopted a hybrid system,[131] such as India), computer programs are not patentable "as such".[132] Features of the computer program,[133] as well as the presence of a device defined in the claim[134] may lend technical character. Moreover, a computer program by itself can be patented if it brings about a further technical effect going beyond the normal physical interactions between the said program and the computer.[135] In the UK, after *Symbian v Comptroller-General of Patents,*[136] the focus is not on the question whether the contribution falls within the excluded subject matter,[137] but on whether the invention makes a technical contribution to the known art, even if the computer program does not bring any novel effect outside of a computer.[138]

v EasyJet Airline Co Ltd (III) [2004] EWHC 1725 (Ch). On the graphic user interface (GUI), see *Bezpečnostní softwarová asociace - Svaz softwarové ochrany v Ministerstvo kultury* [2010] ECR I-13971.

125 Daniel J.M. Attridge, 'Challenging claims! Patenting computer programs in Europe and the USA' (2001) 1 Intellectual Property Quarterly 22.

126 Guido Noto La Diega, 'Software patents and the Internet of Things in Europe, the United States, and India' (2017) 39(3) European Intellectual Property Review 173.

127 As shown by Daehwan Koo, 'Patent and copyright protection of computer programs' (2002) Intellectual Property Quarterly 172, 173, "[p]atent and copyright do not provide optimum protection for software innovations, because they are based on exclusive property rights which impede follow-on small-scale innovations such as software innovations".

128 The search engine of the European Patent Office does not allow to retrieve data of a similar granularity.

129 These data and the following data, including those used in Table 1, were retrieved (and partly calculated) using the US Patent and Trademark Office (USPTO) Patent Full-Text and Image Database, accessed on 9 September 2017. The method is the same used by Allen Clark Zoracki, 'When Is an Algorithm Invented: The Need for a New Paradigm for Evaluating an Algorithm for Intellectual Property Protection' (2005) 15 Alb. L.J. Sci. & Tech. 579, 585, with regards to the patents granted between 1994 and 2003.

130 The faster growth of software patents may be related also to the fact that algorithms can be perceived as mere abstract ideas non-eligible for patent protection. Therefore, some applicants may purposely shy away from using the word 'algorithm' in the specifications.

131 There are mainly three systems for the protection of computer programs. First, one may refer to the double binary of copyright and patent protection, as exemplified by the US approach. Second, there is the hybrid system where alongside copyright, there is a rule excluding computer programs from patentability, but only if claimed 'as such.' This is the system that one finds in Europe. Finally, there is single binary (only copyright) protection. This system is the least common, see the Philippines, which are moving towards the hybrid system. There is a convergence between the double binary and the hybrid systems, with a trend towards a *de facto* generalised double binary.

132 European Patent Convention, art 52(2).

133 T 1173/97 (Computer program product) of 1 July 1998.

134 T 0424/03 (Clipboard formats I/MICROSOFT) of 23 February 2006; T 0258/03 (Auction method/HITACHI) of 21 April 2004.

135 G 3/08 (Referral by the President of the EPO in relation to a point of law ... of 16 October 2009; T 1173/97 (Computerprogrammprodukt) of 1 July 1998.

136 [2008] EWCA Civ 1066, [16] [49] [51 [59].

137 This was the law under *Aerotel v Telco Holdings* [2007] 1 All ER 225, [40].

138 *Shopalotto.com Ltd, Re patent application GB 0017772.5*: PATC 7 November 2005 [2005] EWHC 2416 (Pat), found that in

40 Even though some courts or examiners may consider algorithms as computer programs, they should probably be more precisely seen as mathematical methods. The European Patent Office's Board of Appeals stated that algorithms are mathematical methods, as such deemed to be non-inventions; therefore, a technical character of the algorithm can be recognised only if it serves a technical purpose.[139] The fact that a computer-implemented invention includes an algorithm can make the latter patentable. Indeed, it has been recognised that mathematical algorithms may contribute to the technical character of an invention, inasmuch as they serve a technical purpose.[140] For example, text classification does not qualify as technical purpose.[141] A technical effect may arise either from the provision of data about a technical process, or from the provision of data that is applied directly in a technical process.[142] However, the inclusion of an algorithm in a patent application for a computer-implemented invention does not, in itself, ensure patentability. Indeed, not all efficiency aspects of an algorithm are by definition without relevance for the question of whether the algorithm provides a technical contribution. However, such technical considerations must go beyond merely finding a computer algorithm to carry out some procedure.[143] In the US, legal scholars[144] have focused on how to evidence an improvement in algorithmic technique. It has been suggested to run the algorithm on test problems with known solutions and compare the results with those of algorithms in the prior art, with particular regards to speed, performance, memory usage, and ease of implementation.[145]

41 Unlike copyright, most uses of a computer-implemented invention are prohibited if not authorised and maybe that is why scholars tend to overlook patent exceptions.[146] However, in proceedings for infringement, defendants may avail themselves of the private non-commercial use[147] and experimental use[148] defences. One can qualify for the first immunity even when the resulting information has a commercial benefit, or the subjective intention was not commercial.[149] This is particularly interesting because in the UK there is no private copy exception to copyright.[150] As to the second defence, activities to discover something unknown, to test a hypothesis or to assess whether an invention works are considered as experiments and non-infringing.[151] However, this defence may be of limited use in the context of accessing algorithms, because it cannot be invoked to show that a product works in the way claimed by the maker.[152] Yet, arguably, when accessing the algorithm, the affected individual would have an interest to show that the algorithm-related invention does *not* work in the way claimed by the maker. Thus, this defence could be usefully invoked when an algorithm-related invention is used to take decisions whose rationale one wants to contest.

42 Intellectual property seems to create more problems than solutions to the issue at hand. The route above is weak for at least four reasons. First, the overlap between, if not abuse of, intellectual property rights[153] create a legal black box which is very difficult to open. Second, the application of the study and observation exception presupposes the lawful use of a copy of the software,[154] which is rarely the case in the event of algorithmic decisions. Third, even though the analysed copyright exceptions have been qualified as quasi-rights, there is no precedent

a claim for a lottery game played on the internet, the technical effect did not go beyond the mere loading of a program into a computer.

139 T 1784/06 (Classification method/COMPTEL) of 21 September 2012. In the UK, in *Gale's Application* [1991] RPC 305, it was held that an algorithm used to calculate square roots could not be patented because it lacked any technical character.

140 Ibid. 3.1.1. See also T 2249/13 (Mobile device/TRADE CAPTURE) of 17 October 2014.

141 T 1358/09 (Classification/BDGB ENTERPRISE SOFTWARE) of 21 November 2014; T 1316/09 of 18 December 2012.

142 T 1670/07 (Shopping with mobile device/NOKIA) of 11 July 2013.

143 T 1358/09 (n 142); see G 3/08 (n 136). *HTC Europe Co Ltd v Apple Inc* [2013] EWCA Civ 451, provides a good guidance to understand if computer programs and algorithms are patentable because the invention produces a technical effect that goes beyond the excluded subject matter.

144 Zoracki (n 129) 579.

145 ibid 605.

146 See David Gilat, *Experimental use and patents* (Wiley 1995); Alan J Devlin, 'Restricting experimental use' (2009) 32(2) Harvard Journal of Law and Public Policy 599; Jessica C Lai, 'A right to adequate remuneration for the experimental use exception in patent law: collectively managing our way through the thickets and stacks in research?' (2016) 1 Intellectual Property Quarterly 63.

147 Patents Act 1977, s 60(5)(a).

148 Patents Act 1977, s 60(5)(b).

149 *SKF Laboratories Ltd v Evans Medical Ltd* [1989] FSR 513.

150 See Guido Noto La Diega, 'In Light of the Ends. Copyright Hysteresis and Private Copy Exception after the British Academy of Songwriters, Composers and Authors (BASCA) and Others v Secretary of State for Business, Innovation and Skills Case', in *Studi giuridici europei 2014* (C Franchini ed, Giappichelli 2016) 39.

151 *Monsanto Co. v Stauffer Chemicals Co. and another* [1985] RPC 515 (CA); *Micro-Chemicals et al. v Smith Kline and French Inter-American Corporation* (1971) 25 DLR 78, 89.

152 *Monsanto* (n 152) 542; *Auchinloss v Agricultural and Veterinary Supplies Ltd* [1999] RPC 397, 405.

153 *cf*, more generally, Neil Wilkof and Shamnad Basheer (eds) *Overlapping Intellectual Property Rights* (OUP 2012).

154 Only the 'person having a right to use a copy of a computer program' can avail themselves of the exception (Software Directive, art 5(3).

interpreting said exceptions to open the algorithmic black box. Lastly, it requires considerable skills to open an algorithm by observing and studying the software that implements it. In most cases, there would be the need to ask an expert third party to carry out such activities on behalf of the lawful user of the software. However, applying *SAS Institute*,[155] it is unclear whether said third parties would qualify as lawful users. In the negative, this exception would be of little use in the majority of cases.

43 To add to the complexity, intellectual property will always be balanced with competing interests, such as data protection. As correctly pointed out, for instance, "trade secrecy (...) may make it difficult for data controllers to comply with their obligation of transparent processing."[156] Let us have a look, therefore, at the relevant data protection regime.

D. Algorithmic decision-making and EU data protection

44 The use of algorithms is under the lens of the data protection authorities, especially with regards to profiling. The European Data Protection Supervisor[157] has pointed out that the problem is not profiling as such, but "the lack of meaningful information about the algorithmic logic which develops these profiles and has an effect on the data subject."[158]

45 Under the Data Protection Directive,[159] there is a right not to be subject to a decision which produces legal effects or significantly affects the data subjects, if the decision is based solely on automated processing of data aimed at evaluating certain personal aspects concerning them (e.g. creditworthiness). Moreover, there is a right to know the logic involved in any automated processing of data.[160] Nonetheless, one may be subject to an algorithmic decision in two scenarios.[161] Firstly, in the course of the entering into a contract (or of the performance thereof), provided the request for the entering into the contract (or the performance thereof), lodged by the data subject, has been satisfied or that there are suitable measures to safeguard his legitimate interests (e.g. the data subject could express their viewpoint). For instance, some law firms[162] are using AI-enabled computer programs to assess the merits of personal injury cases and decide, therefore, whether to accept the case or to draft contingency fee agreements. Secondly, and more generally, algorithmic decision-making may be authorised by a law, if there are measures to safeguard the data subject's legitimate interests.[163] Fraud and tax evasion prevention are the typical examples.[164]

46 The rules on algorithmic decision-making have been amended by the GDPR,[165] which is set to come into effect on 25 May 2018, also in the UK, regardless of Brexit.[166] The general principle is that data subjects should not be subject to algorithmic decisions. However, when non-human agents take a decision that has legal effects on the data subject's life "or similarly significantly affects him or her,"[167] the data subject has the rights to obtain human intervention, to express their point of view, as well as to contest the decision.[168] Correspondingly, the data controller

155 *SAS Institute IV* (n 108).

156 Kamarinou (n 76) 23.

157 The European Data Protection Supervisor is the EU data protection authority. They inter alia ensure the protection of personal data and privacy when EU institutions and bodies process personal data. See Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data [2001] OJ L 8/1.

158 European Data Protection Supervisor, 'Recommendations on the EU's options for data protection reform' (2015/C 301/01), para 3.1.

159 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (Data Protection Directive), art 15.

160 Data Protection Directive, art 12(a); recital 41. For a national implementation, see the UK Data Protection Act 1998, s 7(1)(d).

161 Data Protection Directive, art 15(2).

162 See Jane Croft, 'Legal firms unleash office automatons' (*The Financial Times*, 16 May 2016), https://www.ft.com/content/19807d3e-1765-11e6-9d98-00386a18e39d.

163 In the UK, the Secretary of State may prescribe in which circumstances (apart from a contract) an algorithmic decision may be exempt from the said rules (Data Protection Act 1998, s 12(5)(b)).

164 Information Commissioner's Office, *Overview of the General Data Protection Regulation (GDPR)* (ICO 2017) 27. The same example can be found in the GDPR, recital 71.

165 The underlying principle is the same, that is that "fully automated assessments of a person's character should not form the sole basis of decisions that significantly impinge upon the person's interests" (Lee Bygrave, 'Automated Profiling, Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling' (2001) 17(1) Computer Law & Security Review 17, 21) as suggested by Kamarinou (n 76) 8.

166 In the time between 25 May 2018 and 29 March 2019 the rules about algorithmic decision-making will be those resulting from a combination of GDPR and the Data Protection Act, after 29 March 2019 it is likely that only the Data Protection Act as amended will be in force. *cf* The Rt Hon Karen Bradley MP, Culture, Media and Sport Committee, *Oral evidence: Responsibilities of the Secretary of State for Culture, Media and Sport*, HC 764 (24 October 2016).

167 GDPR, art 22.

168 As to the latter, it would seem to us that this right as enshrined in art 22 of the GDPR is the same as the right to "challenge the decision" under recital 71. *Contra*, see Kuner (n 76) 2, who observe that even though the recital is not

shall provide "meaningful information about the logic involved"[169] in the algorithmic decision. It is likely that the national implementing measures of the Data Protection Directive will be amended or replaced to recognise a stronger protection to data subjects against algorithmic decisions.[170]

I. The general prohibition on solely automated decisions with a significant effect

47 Let us start with the provisions directly dealing with algorithmic decision-making;[171] it is open to debate whether they constitute a considerable step forward. The main right available to the data subject is the right not to be subject to a solely automated decision with legal effects or similarly significantly affecting them.[172] This can be interpreted as a general prohibition to make algorithmic decisions using personal data, or as a mere right to be oppose (after being informed about) the algorithmic decision.[173] In the UK, data subjects can require that no solely algorithmic decision be taken against them. However, if no such notice has effect and the decision is taken, the data controller has 21 days to give a written notice explaining the steps that they will take to comply with the data subject request.[174] Positively, in issuing some guidelines on algorithmic decision-making, the Article 29 Working Party[175] recommends treating this right as a general prohibition.[176] Regrettably, however, the only amendment introduced by the Data Protection Bill with regards to algorithmic decisions concerns the safeguarding measures that controllers should take when availing themselves of the consent-based exception.[177] Arguably, by refusing the "general prohibition" approach, the UK will not comply with the GDPR, with practical consequences for instance in terms of the legality of the EU-UK data transfers. If this provision expresses a core data protection principle,[178] a partial compliance may cause the EU to deem the UK protection of personal data inadequate, hence hindering cross-border data flows.[179]

48 Looking at the core of art 22, there are two main differences between the Data Protection Directive and the GDPR.

49 First, in the new provision there is an express reference to profiling as an example of automated processing. This brings clarity in a field currently perceived as particularly relevant, but it risks

binding, it "may embolden regulators and courts to try to compel data controllers to provide explanations of specific outcomes in particular cases, and not merely 'meaningful information' about 'logic'".

169 GDPR, arts 13(2)(f) and 14(2)(g).

170 In August 2017, the UK government announced a new Data Protection Bill, where "individuals will have greater say in decisions that are made about them based on automated processing. Where decisions are based on solely automated processing individuals can request that processing is reviewed by a person rather than a machine." (UK Department for Digital, Culture Media & Sport, A new data protection bill: Our planned reforms, The UK Government (7 August 2017), <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/635900/2017-08-07_DP_Bill_-_Statement_of_Intent.pdf> accessed 3 March 2018). The Bill was introduced in the House of Lords on 13 September 2017 and it passed second reading at the House of Commons on 5 March 2018. Members of Parliament are considering the Bill in a Public Bill Committee, which is set to finish by 27 March 2018. See below for the analysis of s 14 of the Bill (as brought from the Lords), regarding algorithmic decision-making authorised by the law. It must be said that the fact that the only relevant provision in the Bill regards the limited issue of the algorithmic decisions authorised by law may be seen as a missed opportunity to thoroughly review the regime laid out in s 12 of the Data Protection Act 1998.

171 The focus of this section is on the rules regarding algorithmic decision-making. These do not apply to profiling *per se* if it is not followed by an algorithmic decision producing legal effects or similarly significantly affecting the data subject. For more information on the rules about profiling, regardless of whether or not it is followed by an algorithmic decision, please see ibid 17-25.

172 GDPR, art 22(1).

173 For the first interpretation, in favour of a general prohibition of algorithmic decisions, see the French *Loi n° 78-17* of 6 January 1978 *relative à l'informatique, aux fichiers et aux libertés*, art 10.

174 Data Protection Act 1998, s 12(3).

175 The Article 29 Working Party is an advisory body set up under the Data Protection Directive, art 29. It is composed by representatives from the Member States' data protection authorities, the European Data Protection Supervisor and the European Commission. The GDPR will replace it with the European Data Protection Board.

176 Article 29 Working Party (n 9) 12.

177 Other Member States are adopting implementing measures that are overlooking algorithmic decision-making. For instance, on 21 March 2018, the Italian Cabinet (*Consiglio dei Ministri*) adopted the draft decree implementing the GDPR (*Disposizioni per l'adeguamento della normativa nazionale alle disposizioni del Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati, e che abroga la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati*), hereinafter 'Draft GDPR implementing decree'). The Draft GDPR implementing decree does not provide anything on the matter. Unlike the UK, however, the GDPR will be directly applicable and, therefore, Italian data controllers will be bound directly by art 22 of the GDPR.

178 This is the view expressed by Bygrave (n 165) 22 about the similar provision in the Data Protection Directive. His observation is all the more true with regards to the GDPR for at least two reasons. First, because algorithmic decisions have become more common and more intrusive. Second, because the GDPR strengthens the relevant regime, thus confirming the importance of the provision.

179 GDPR, art 45.

narrowing the interpretation of the provision thus excluding forms of algorithmic decision-making which do not include profiling. Therefore, it is positive that the Article 29 Working Party has observed that "(a)utomated decisions can be made with or without profiling; profiling can take place without making automated decisions."[180]

50 Second, and most importantly, one has the said right only if the decision produces legal effects concerning one "or *similarly* significantly affects him or her."[181] This addition goes in the opposite direction to the one taken when the draft GDPR was first published and it had been suggested that art 22 should cover not only decisions producing legal effects or which significantly affect data subjects, but also the "collection of data for the purpose of profiling and the creation of profiles as such."[182]

51 Now, "legal effect" is quite straightforward, including all the scenarios where a decision affects a person's rights based on laws or contracts.[183] In turn, "similarly" may narrow the scope of the provision, if compared with the previous wording, where no reference to this adverb was made. Indeed, it may be seen as meaning that one does not have the right to object to algorithmic decision-making if the effect is not similar to a legal effect[184] (e.g. significant distress or missed professional opportunities as a consequence of being permanently banned from a popular social network).[185] If this interpretation were followed, broader national implementations may need to be reviewed accordingly. For instance, the UK refers to decisions take for "the purpose of evaluating matters relating to him such as, for example, his performance at work, his creditworthiness, his reliability or his conduct"[186]. The Information Commissioner's Office accepts that it is hard to explain what "significant effect" means, but it suggests that it refers to "some consequence that is more than trivial and potentially has an unfavourable outcome."[187] Businesses have been asking for more detailed guidance[188] and this has partly arrived with the Article 29 Working Party's guidelines that indicated that "similarly" means that "the threshold for *significance* must be similar."[189] Therefore, in order for a decision to fall within the scope of art 22, it must not necessarily be a quasi-legal effect in terms of content, being sufficient a decision which profoundly affects the individual as much as a decision affecting her or his rights would. Adding details to the UK attempt of definition, the EU advisory body point out that a similarly significant effect must be "more than trivial and must be sufficiently great or important to be worthy of attention."[190] The concept is broad enough to encompass a vast number of scenarios, from e-recruiting to online behavioural advertising, especially if intrusive and targeted to vulnerable groups,[191] as well as consumer manipulation.[192]

52 Even before understanding what 'legal' means, one should clarify what a 'decision' is. It has been suggested that this could include "an interim or individual step taken during the automated processing."[193] It would seem, however, that only rarely interim measures and individual steps will qualify for the application of art 22 of the GDPR, because the provision requires a decision with legal effect or "similarly significant."

53 Some aspects of this regime are not clear yet. For instance, it is open to debate what *solely* automated means. In the past, it was relatively easy to understand what 'solely' meant. There was a limited number of organisations taking significant algorithmic decisions and the technologies used were

180 ibid 8.

181 GDPR, art 22(1).

182 Article 29 Working Party, 'Advice paper on essential elements of a definition and a provision on profiling within the EU General Data Protection Regulation' (13 May 2013), para 2(a).

183 *cf* ibid 10.

184 This would constitute a weakening of many national implementing regimes. For instance, in the UK the Data Protection Act 1998 refers generally to decisions which significantly affect the individual (s 12(1)).

185 *cf* Jilian York, 'Getting banned from Facebook can have unexpected and professionally devastating consequences' (Quartz, 31 March 2016) <https://qz.com/651001/getting-banned-from-facebook-can-have-unexpected-and-professionally-devastating-consequences/> accessed 1 March 2018.

186 Data Protection Act 1998, s 12(1).

187 Information Commissioner's Office, *Feedback request – profiling and automated decision-making* (ICO 2017) 19.

188 The digital technology industry in Europe would welcome such guidance. For instance, 'Input on Automated Individual Decision Making & Data Breach Notification' (DigitalEurope, 5 April 2017) 3 <http://www.digitaleurope.org/DesktopModules/Bring2mind/DMX/Download.aspx?Command=Core_Download&EntryId=2390&language=en-US&PortalId=0&TabId=353> accessed 1 March 2018, "would appreciate (...) clarification in the future guidance on how companies should interpret these two cumulative conditions as well as examples of such effects in different sectors".

189 Article 29 Working Party (n 9) 10.

190 ibid 10.

191 *cf* Guido Noto La Diega, 'Some considerations on intelligent online behavioural advertising' (2017) 66-67 Revue du droit des technologies de l'information 53.

192 Artificial intelligence is increasingly used to predict consumers' behaviour in order to lock them in by means of addiction. Evidence has been recently uncovered about such manipulating practices in the gambling industry. See Mattha Busby, 'Revealed: how bookies use AI to keep gamblers hooked' (*The Guardian*, 30 April 2014) <https://www.theguardian.com/technology/2018/apr/30/bookies-using-ai-to-keep-gamblers-hooked-insiders-say> accessed 2 May 2018.

193 ibid 12.

quite rudimental; therefore, reviewing the machine-generated data was relatively straightforward and once a human being reviewed the data, the decision was no longer solely automated.[194] In light of increasingly complex (and accordingly opaque) algorithmic techniques and of the ubiquitous nature of the phenomenon of algorithmic decisions, that approach should be abandoned. To what extent is the human intervention meaningful vis-á-vis black-box decisions?

54 The UK Information Commissioner's Office recently requested feedback on some points of the GDPR,[195] and they have suggested that 'solely' should "cover those automated decision-making processes where a human exercises no real influence on the outcome of the decision, for example where the result of the profiling or process is not assessed by a person before being formalised as a decision."[196] The risk of this interpretation is that it is not always easy - especially from the data subject's perspective - which role the human being played in the decision (was the human being a passive operator? Which discretion did they have while assessing the result?). Moreover, "it may not be feasible for a human to conduct a meaningful review of a process that may have involved third-party data and algorithms (which may contain trade secrets), prelearned models, or inherently opaque machine learning technique."[197] Therefore, it would seem more appropriate to recognise the right not to be subject to an algorithmic decision every time that there is not a human being clearly taking the final decision.[198] It would seem that the Article 29 Working Party hold similar views when they state that a decision is not wholly automated when alongside an automated profile, there is "additional meaningful intervention carried out by humans before any decision is applied to an individual."[199] However, there is still a lack of clarity. Indeed, in order to clarify when art 22 GDPR applies or not, the Article 29 Working Party makes the following examples. If a human decides whether to agree the loan based on a profile produced by purely automated means, then art 22 will not apply. In turn, if an algorithm decides whether the loan is agreed and the decision is automatically delivered to the individual, without any meaningful human input, then art 22 will apply. The point is that there is a substantial grey area here. For instance, it is unclear whether art 22 applies when the algorithmic system takes the decision, but a human being reviews it. Arguably, the human review could qualify as "meaningful human input", but this will have to be assessed on a case-by-case basis.

55 Even more importantly, controllers should refrain from "fabricating human involvement"[200] with the purpose of sidestepping art 22; this provision will apply every time that there is not meaningful and genuine human intervention, for instance in the form of actual oversight by a person with "authority and competence to change the decision."[201] It is important to stress that the GDPR applies to every automated profiling carried out on personal data to evaluate a natural person's personal aspects, not only to the 'solely' automated one, which means that the general GDPR rules and standards will apply to profiling even when a human being plays a substantial role in the creation of the relevant profile.[202]

II. Three exceptions: contract, consent, law

56 Even though "as a rule, there is a prohibition on fully automated individual decision-making (...) that has a legal or similarly significant effect,"[203] this rule has some exceptions. The GDPR has innovated the systems of the exceptions not only by adding a consent-based exception, but also by clarifying the scope of the pre-existing ones. It is unfortunate that the UK Data Protection Bill[204] is missing out on this opportunity. Indeed, the only innovation that it is being introduced regards algorithmic decisions authorised by law. The UK will keep allowing such decisions in circumstances prescribed by the Secretary of State, in relation to a contract, when authorised or required by or under any enactment, effect of the decision is to grant a request of the data subject, or when steps have been taken to safeguard the legitimate interests of the data subject.

194 This is still the approach that one can find in Information Commissioner's Office, 'Guide to Data Protection' (*ICO*, 11 May 2016) <https://www.legislation.gov.uk/ukpga/1998/29/pdfs/ukpgaod_19980029_en.pdf> accessed 18 March 2018.

195 Information Commissioner's Office (n 186).

196 ibid 19.

197 Kuner (n 76) 2.

198 Along the same lines, with regards to the Data Protection Directive, it has been noted that the regime will operate every time that there is not a human being exercising "real influence on the outcome of a particular decision-making process" (Bygrave (n 165) 20).

199 Article 29 Working Party (n 9) 8.

200 ibid 10. As an example, they observe that "if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing." (ibid 10).

201 ibid 10.

202 ibid 6. For instance, profiling is rarely transparent. However, the controller must provide data subjects with concise, transparent, intelligible and easily accessible information about the processing of their personal data (GDPR, art 12(1)).

203 ibid 9.

204 Data Protection Bill [HL] 2017-19, s 14.

No consent-based exception is provided. Unlike the interpretation of the right not to be subject to an algorithmic decision as a general prohibition, the lack of implementation of the consent-based exception is unlikely to endanger the cross-border data transfers with the EU. Indeed, a lack thereof might ensure a stronger protection of personal data. In turn, the broad wording of the contractual exception may be more problematic.[205]

57 Art 22 brings clarity to the scenario regarding the entering and performance of the contract by simplifying the language and restricting the contractual exception to the instances when the algorithmic decision-making is *necessary* to enter into a contract or for its performance.[206] One may argue, going back to the example of the contingency fee agreements, that in that scenario the algorithmic decision would not be necessary and, thus, it would not fall within the scope of this exception. Following the European Data Protection Supervisor's approach, if a less privacy-intrusive method is available, then the algorithmic decision is not necessary and, therefore, it is not allowed.[207]

58 In turn, the new exception based on the data subject's explicit consent[208] is problematic. Consent is explicit when there is "an express statement rather than some other affirmative action."[209] Indeed, given the imbalance of bargaining power that characterises many transactions, one should not be surprised if, for instance, a bank could force a potential client requesting a loan to consent to a decision taken by an algorithm. The exception based a law authorising the decision while laying down measures to safeguard the data subject's legitimate interest[210] now includes a reference to the data subject's rights and freedoms and to both EU and national law. These changes are nugatory. Firstly, based on an *a minore ad maius* argument, it is obvious that if the decision shall respect the legitimate interests of the data subject, it shall do so also with regards to the more relevant rights and freedoms. Secondly, while the reference to national laws is a truism, the one to EU law cannot be interpreted as a power to legislate beyond what already provided by the treaties. However, the growth of artificial intelligence (AI) may have an impact on the analysed regime. Not only because, generally, AI does not always make it feasible to access the rationale of algorithmic decisions. With specific regards to the consent-based exception, it is fair to wonder, "how can informed consent be obtained in relation to a process that may be inherently non-transparent (a 'black box')."[211]

59 The third exception regards national and EU laws authorising algorithmic decisions.[212] Regrettably, the Article 29 Working Party do not provide any guidance on the matter. Whereas recital 71 refers only to fraud, tax evasion, and reliability of the service, it would seem that EU and national authorities may allow algorithmic decisions for a potentially infinite number of purposes. Indeed, recital 73 provides that EU and national laws can impose restrictions concerning "decisions based on profiling" in inter alia order to prevent or react to breaches of ethics for regulated professions or for the keeping of public registers kept for reasons of public interest. Therefore, for instance, a Member State could allow algorithmic decisions to disbar a barrister who behaved unethically. Nor are there limits to which kind of public registers a state may keep, for instance for surveillance purposes.[213] One should not think, however, that if a law authorises the algorithmic-decision making in a specific field, say fraud, data protection legislation can be eluded altogether. Alongside the rights to access, the information rights and right to a human judge, data controllers will still have to comply with all the other data protection principles, including accountability.[214] The Data Protection Directive required the laws authorising algorithmic decisions to safeguard only the data subjects' legitimate interests and not also their rights

205 The Data Protection Act 1998 enables data controllers to make algorithmic decisions in the course of steps taken for the purpose of considering whether to enter into a contract, with a view to entering into such a contract, or in the course of performing such a contract, or if the decision is authorised or required by or under any enactment (Data Protection Act 1998, s 12(6)).

206 GDPR, art 22(2)(a).

207 *cf* European Data Protection Supervisor, 'Assessing the necessity of measures that limit the fundamental right to the protection of personal data. A Toolkit' (European Data Protection Supervisor, 11 April 2017), <https://edps.europa.eu/sites/edp/files/publication/17-04-11_necessity_toolkit_en_0.pdf> accessed 9 March 2018.

208 GDPR, art 22(2)(c).

209 Article 29 Working Party (n 9) 13. These guidelines do not provide sufficient clarity as to how to ensure explicit consent. The matter will be addressed in the forthcoming consent guidelines.

210 GDPR, art 22(2)(b).

211 Kuner (n 76) 1.

212 GDPR, art 22(2)(b).

213 Nonetheless, the restrictions should be necessary and proportionate in a democratic society to safeguard public security and in compliance with Charter and in the European Convention for the Protection of Human Rights and Fundamental Freedoms.

214 The GDPR provides that the algorithmic decision-making for purposes authorised by EU or national law should be "conducted in accordance with the regulations, standards and recommendations of Union institutions or national oversight bodies." (recital 71). Even though it may be interpreted as referring only to fraud and tax evasion, it would be absurd to exclude other purposes specifically authorised by the law (the reference is illustrative, not exhaustive). It is submitted that the data protection authorities should be deemed as oversight bodies and the data protection laws should still apply even when an algorithmic decision is allowed.

and freedoms. Moreover, it did not specify which laws could authorise algorithmic decisions. The GDPR, in turn, now includes a reference to the data subject's rights and freedoms and it clarifies that both EU and national laws can authorise algorithmic decisions. Arguably, these changes are nugatory. Firstly, based on an *a minore ad maius* argument, it is obvious that if the decision should respect the legitimate interests, all the more it should do so with rights and freedoms. Secondly, the clarification that national law can be a legal basis is a truism. So is the one about EU law, which should not be interpreted as a power to legislate beyond what already provided by the treaties.

60 The UK Data Protection Bill[215] provides more detail as to the procedure to follow when an algorithmic decision falls under the third exception. Indeed, the controller must, as soon as reasonably practicable, notify the data subject in writing that a decision has been taken based solely on automated processing. Correspondingly, the data subject may, before the end of the period of 21 days, beginning with receipt of the notification, request the controller to reconsider the decision, or take a new decision that is not based solely on automated processing. The provision goes on to point out what the controller must do if such request is made. The procedure is the same that the Data Protection Act currently provides for non-exempt decisions, but interestingly the new regime is more protective of the data subject if compared to the previous one. Indeed, currently the data controller's notice must only indicate the steps the controller intends to take to comply with the request. This information must be notified before the end of the period of 21 days beginning with receipt of the request. On top of this, the Data Protection Bill provides that when the law authorises an algorithmic decision, the data controller shall consider the request, comply with it, and inform the data subject of the steps taken to comply, and of the outcome of complying with the request. The wording suggests that data controllers have some discretion in complying. However, the discretion regards how to comply, not whether to comply. The only reason why a denial could be allowed would be if the algorithmic decision was not taken solely on the basis of automated processing, if the decision does not significantly affect the data subject, or if it is impossible to identify the data subject.[216] If the data controller violated the limits of its discretion, the data subject may appeal the decision judicially.

61 Interpreters will need to avoid a visible inconsistency in the new UK regime on algorithmic decision-making. Namely, it is not rational to give the data subject a weaker protection when a non-exempt decision is at issue, if compared to a decision authorised by the law.

62 One may observe a departure of UK data protection law from the GDPR. In the UK, there is a three-layered system. As a rule, data subjects must be informed of non-exempt algorithmic decisions and can request that no such decision be taken. If no request has effect, they still have a right to be informed and to request a reconsideration or a human decision. Reconsideration and the right to a human judge, after the Data Protection Bill is enacted, will apply also to the algorithmic decisions authorised by law. Obviously, no right to pre-empt such a decision would apply. Thirdly, data subjects have no rights regarding the other exempt decisions.[217] This may raise concerns in terms of adequacy of the protection of personal data in the UK in the context of cross-border data transfer with the EU. Since consent is not one of the exceptions, the rights of the first layer will apply. In the EU, in turn, there is a clearer and stronger model. The rule is the general prohibition to take solely algorithmic decisions. There are only three justifications that can be used to make some decisions, but all of them are accompanied by strong safeguards for the data subject.

63 Lastly, it is not entirely clear if the list of exceptions (contract, consent, law) is exhaustive. A recital[218] refers to algorithmic decision-making for the purpose of ensuring the security and reliability of a service provided by the controller. However, this should not be interpreted as a fourth exception or as proof of the non-exhaustive character of the list of exceptions. It is plausible, indeed, that this is only an example of a purpose for which national and EU laws can authorise the said decision-making.[219]

III. Measures to safeguard the data subjects' rights, freedoms, and legitimate interests

64 The main commendable innovation in the GDPR regards the measures to safeguard the data subject's rights, freedoms, and legitimate interests affected by

215 Data Protection Bill [HL] 2017-19, s 14.

216 The GDPR is very clear in stressing that the data controller "shall not refuse to act on the request of the data subject for exercising his or her rights under Articles 15 to 22, unless the controller demonstrates that it is not in a position to identify the data subject" (art 12(2).

217 Under the Data Protection Act, s 12(4), the data subject's request not to take solely algorithmic decisions does not have effect in relation to an exempt decision; and nothing in s 12(2), regarding the data controller's notice, applies to an exempt decision.

218 GDPR, recital 71.

219 See Article 29 Working Party (n 9) 12.

an algorithmic decision.[220]

65 First, now these measures refer also to the contractual and consent-based exceptions. Second, they are no longer limited to the right to express one's viewpoint. The provision shall be interpreted as the right to obtain human intervention on the part of the controller and the right to contest the decision. Therefore, if there is a law authorising algorithmic decision making,[221] if this is necessary for a contract, or if there is the data subject's explicit consent, a data controller may use algorithms to take decisions having legal effects or similarly affecting the data subject. However, data controllers shall put in place a procedure to appeal the decision with meaningful oversight by a human being that shall ensure an effective right of defence to the data subject.[222]

66 This is a major victory for those who think that human decision-making is still better than the automated one.[223] However, it is unclear which steps the data controller should take once the data subjects avail themselves of the analysed remedy. The Article 29 Working Party further clarify that the review must be carried out by a human being with appropriate authority and capability to change the decision and who shall thoroughly assess "all the relevant data, including any additional information provided by the data subject."[224]

220 GDPR, art 22(3).

221 The wording of the provision is not crystal clear. Indeed, art 22(2)(b) applies the said measures to the algorithmic decision-making authorised by the law. Then, the following paragraph extends these measures to the other two exceptions and it specifies that they include "at least" the right to human intervention, to express the viewpoint, and to contest the decision. It may be argued, therefore, that when the law authorises algorithmic decision-making, the mere right to express one's viewpoint (as provided under the old regime) would be sufficient. However, this would seem to go against the overall purpose of the GDPR and of art 22. Moreover, the express reference to "at least" is likely to mean that those three rights are the minimum core of the measures that safeguard the data subject's rights, freedoms, and legitimate interests. Furthermore, recital 71 suggests that these measures should be put in place "[i]n any case".

222 Obviously, if such a system is not in place or if the data subject is not satisfied, the usual judicial remedies will be available.

223 A slightly different perspective is taken by Kamarinou (n 76) 22, who observe that "it may already in some contexts make sense to replace the current model, whereby individuals can appeal to a human against a machine decision, with the reverse model whereby individuals would have a right to appeal to a machine against a decision made by a human".

224 Article 29 Working Party (n 9) 15.

IV. Transparency obligations: a right to explanation?

67 Moving onto the transparency obligations, these are nearly entirely new,[225] given that under the Data Protection Directive there was only the right to access, which included the logic involved in the algorithmic decision.[226] Innovatively, the processing is not deemed fair and transparent, if the controller does not - at the time when personal data is obtained from the data subject - provide specific information on three matters.[227] First, controllers must disclose the existence of algorithmic decision-making. Second, they need to inform the data subject about the logic involved. Third, the algorithm must be opened in order to provide "meaningful information about [...] the significance and the envisaged consequences of such processing for the data subject."[228] The same right applies when the data was not obtained from the data subject, who has the right to be informed within a reasonable timeframe[229] (at the latest within one month),[230] at the time of the first communication with the data subject,[231] or when the data is fist disclosed to a third party.[232] Data controllers who merely make the information available, without actively bringing it to the data subject's attention, do not meet their transparency obligations. On top of the obligation to inform, there is the right of access, which again regards the existence of the algorithmic decision-making itself and meaningful information about the logic, the significance, and the consequences.[233]

68 One should welcome positively the obligation to provide (and the right to access) meaningful information and the reference to the envisaged consequences and significance of the decision. While

225 They are new at an EU level, but not necessarily at the national one. For instance, the Data Protection Act 1998 provides the controller's obligation to notify the data subject that the decision was algorithmic (s 12(2)(a)), unless the data subject already required that the decision is not taken based solely on automated processing (s 12(1)-(2)).

226 Data Protection Directive, art 12(a).

227 Information rights exist under the GDPR also when there is no algorithmic decision significantly affecting a data subject. See the principles of fair and transparent processing and arts 13 and 14 of the GDPR. According to Article 29 Working Party (n 9) 13 considers as "good practice to provide the above information whether or not the processing falls within the narrow Article 22(1) definition."

228 GDPR, art 13(2)(f).

229 This is similar to the UK provision, which refers to "as soon as reasonably practicable" (Data Protection Act 1998, s 12(2)(a)).

230 GDPR, art 14(3)(a).

231 GDPR, art 14(3)(b).

232 GDPR, art 14(3)(c).

233 GDPR, art 15(1)(h).

"envisaged" suggests that information must be provided "about intended or future processing,"[234] it would seem that "significance" requires real, tangible examples of how the decision may affect the data subject.[235]

69 Generally speaking, such meaningful information is what the data subject, who normally will not be a computer scientist, is likely to be interested in. Therefore, for instance, a technical document which includes the algorithm used and the mere explanation of the logic in mathematical terms will not in itself meet the legal requirement. Arguably, this should be interpreted as the disclosure of the algorithm with an explanation in non-technical terms of the rationale of the decision and criteria relied upon.[236] Regrettably, the Article 29 Working Party[237] do not consider the disclosure of the algorithm as necessary under the said transparency obligations. However, in order to have a full picture, the data subject has a legitimate interest in asking an expert to analyse the algorithm in order to better challenge the decision. A different interpretation would not comply with right to an effective remedy[238] and to a fair trial[239] under the Charter of Fundamental Rights of the EU and the European Convention of Human Rights.

70 Obviously, it may be the case that, due to the characteristics of artificial intelligence alone, it could be impossible to explain an algorithmic process "in a way that is intelligible to a data subject."[240] However, the data controller should make any reasonable effort to adequately inform the data subject.

71 Scholars have recently criticised the provision because it would entail a right to be informed, but no right to explanation.[241] Others,[242] conversely, have pointed out that Articles 15 and 22 should have a wide interpretation that might prove adequate to cope with the transparency challenge; they propose a legibility stress test for the data controller.

72 To overcome this issue, those who exclude that a right to explanation is provided by the GDPR make a number of recommendations to improve transparency and accountability of algorithmic decision-making, including a trusted third-party regulatory or supervisory body that can investigate algorithmic decisions if one feels that they have been discriminated against. Whereas the idea of an AI watchdog can be a positive one, this paper argues that the information rights provided with regards to algorithmic decision-making – which include a reference to the significance and consequences of the decision – can be interpreted as meaning a right to explanation.[243] Denying it would mean playing down the great potential of legal interpretation. A counterargument could be that the wording 'right to obtain information' can be found in recital 71, but not in art 22; this placement in a non-binding part of the Regulation (a recital) has been seen as "a purposeful change deliberated in trilogue."[244] However, the pivotal role of recitals in interpreting the provisions of an EU act has been expressly recognised by the Commission.[245] The reference to the right of explanation in the recital shall be, therefore, used to properly construe art 22 to reflect the context of the provision and the overall purpose of the GDPR, that is increasing the protection of the data subjects' rights. Hence, even though applying the literal rule, art 22 would not contain a right to explanation, a purposive approach and a correct valorisation of the role of recitals make it clear that data subjects are entitled to such a right. In addition, the data controller is expressly required to provide "concise, transparent, intelligible and easily accessible form, using clear and plain language."[246]

234 Article 29 Working Party (n 9) 14.

235 ibid 14.

236 ibid.

237 This is the interpretation given to recital 60 of the GDPR by Article 29 Working Party (n 9) 13.

238 Charter of Fundamental Rights of the European Union, art 47(1); European Convention on Human Rights, art 13.

239 Charter of Fundamental Rights of the European Union, art 47(2); European Convention on Human Rights, art 6.

240 Kuner (n 76) 1, who suggest that a "high-level, non-technical, description of the decision-making process is more likely to be meaningful" (ibid 2).

241 Sandra Wachter et al., 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7(2) International Data Privacy Law 76. For a similar somehow pessimistic take, see Lilian Edwards and Michael Veale, 'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for' (2017) 16(1) Duke Law & Technology Review 18.

242 Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7(4) International Data Privacy Law 243. On the optimistic front, see also Julia Powles and Hal Hodson, 'Google DeepMind and healthcare in an age of algorithms' (2017) 7 Health Technol 351. Between the two poles, see e.g. Tal Zarsky, 'The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41(1) Science, Technology, & Human Values 118; Mireille Hildebrandt, 'The New Imbroglio - Living with Machine Algorithms', in Liisa Janssens (ed) *The Art of Ethics in the Information Society* (Amsterdam University Press 2016).

243 There is the risk, however, that the courts will interpret the analysed provisions in a narrow way, focusing on the weaknesses of the new regime.

244 Wachter (n 238) 96.

245 Roberto Baratta, 'Complexity of EU law in the domestic implementing process' (2014) 19th Quality of legislation seminar "EU legislative drafting: Views from those applying EU law in the Member States" 4 <http://ec.europa.eu/dgs/legal_service/seminars/20140703_baratta_speech.pdf> accessed 1 March 2018.

246 GDPR, art 12(1).

73 Lastly and commendably, the GDPR details the timescale and procedure to provide information.[247] In particular, the information should be provided without undue delay and in any event[248] within one month of receipt of the request. The information must be in electronic form to reflect the form of the request, unless the data subject requests otherwise.

74 Obviously, the problems with the black boxes remain, no matter how broad the interpretation given to the transparency obligations is. Therefore, the transparency obligations may not be fully effective "in cases where a machine learning process involves multiple data sources, dynamic development, and elements that are opaque, whether for technological or proprietary reasons."[249]

V. Algorithmic decisions with sensitive personal data

75 Another positive new provision regards sensitive personal data (e.g. data on health or sexuality). Artificial intelligence increasingly relies on this kind of data. One need only think that deep neural networks have been recently used to infer the sexual orientation of people from their faces.[250] Indeed, in principle, algorithmic decisions shall not be based on sensitive personal data.[251] For instance, an employer may not let an algorithm decide whether to fire an employee using health data. However, this data may be used with the data subject's explicit consent or in the interest of public health, provided that measures to safeguard the data subject's rights, freedoms, and legitimate interests are in place. Even though ideally it would have been preferable not to have another consent-based exception, unlike the homologous exception regarding non-sensitive personal data, here it is provided that EU or national laws can decide that the prohibition to process sensitive data "may not be lifted by the data subject."[252]

247 GDPR, art 12(3).

248 If the data controller proves that more time is necessary to respond because the request is very complex and there is a high number of requests, there may be an extension by two further months. See GDPR, art 12(3).

249 Kuner (n 76) 2.

250 Yilun Wang and Michal Kosinski, 'Deep neural networks are more accurate than humans at detecting sexual orientation from facial images' (OSFHome, 15 February 2017) <https://osf.io/zn79k/> accessed 1 March 2018 (forthcoming in Journal of Personality and Social Psychology).

251 GDPR, art 22(4).

252 GDPR, art 9(2)(a).

VI. Data Protection Impact Assessments for algorithmic decisions

76 Lastly, one of the main innovations of the GDPR is the data protection impact assessment (DPIA).[253] These impact assessments are tools for organisations to manage data protection hazards, a form of a form of 'meta-regulation' whereby "state efforts to make corporations responsible and accountable for their own efforts to self-regulate."[254] In this field, DPIAs are "a way of showing that suitable measures have been put in place to address those risks (associated to algorithmic decision-making) and demonstrate compliance with the GDPR."[255] It is commendable that DPIAs are mandatory when a systematic and extensive evaluation of personal aspects is based on automated processing, and on which decisions are based that produce legal effects or similarly significantly affect a natural person.[256] Commendably, DPIAs are required both when the decision is wholly automated and when there is human intervention, not only when it is solely based on automated processing.[257]

VII. Can children be subject to algorithmic decisions?

77 An example of poor drafting regards the algorithmic decision-making concerning children. Hidden in a long recital, one finds the obscure sentence "[s]uch measure should not concern a child."[258]

78 Naturally, one would think that children cannot be subject to algorithmic decisions. However, the sentence follows the one that regards the measures to safeguard the data subject's rights, freedoms, and legitimate interests. Therefore, it may be interpreted as meaning that these safeguarding measures do not apply to children, who could nonetheless be

253 See Article 29 Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679' (European Commission, 4 April 2017), <http://ec.europa.eu/newsroom/document.cfm?doc_id=44137> accessed 7 March 2018.

254 This is the theory of Reuben Binns, 'Data protection impact assessments: a meta-regulatory approach' (2017) 7(1) International Data Privacy Law 22, 23. The notion of meta-regulation was developed by C Parker, 'Meta-regulation: Legal Accountability for Corporate Social Responsibility' in D McBarnet, A Voiculescu and T Campbell (eds), *The New Corporate Accountability: Corporate Social Responsibility and the Law* (CUP 2007) 29.

255 Article 29 Working Party (n 9) 27.

256 GDPR, art 35(3)(a).

257 Article 29 Working Party (n 9) 27.

258 GDPR, recital 71.

subject to algorithmic decisions. This is obviously against the purpose of the GDPR, which provides an advanced protection to children. The doctrine of *noscitur a sociis* would lead to absurd consequences; therefore, a purposive approach should prevail. Thus, children should never be subject to algorithmic decision-making.

79 Regrettably, the Article 29 Working Party does not see this provision as an absolute prohibition, since the wording of the recital is not reflected in art 22. However, they recommend that "wherever possible, controllers should not rely upon the exceptions in art 22(2) to justify"[259] algorithmic decision-making affecting children. Nonetheless, such decisions may be necessary for instance to protect the children's welfare, in which case data controllers may resort to the exceptions. Positively, in turn, it is suggested that 'legal effect' and 'similarly significant effect' be interpreted broadly, because "solely automated decision making which influences a child's choices and behaviour could potentially have a legal or similarly significant effect on them, depending upon the nature of the choices and behaviours in question."[260] Similarly, organisations must put in place safeguards tailored to the specific needs and features of the child.[261]

VIII. Collective algorithmic decisions

80 It is unclear, then, what happens to collective algorithmic decisions (e.g. to charge a higher rate of car insurance to the citizens associated to a particular postcode). Indeed, it has been questioned "whether data subjects are protected against decisions that have significant effects on them but are based on group profiling."[262] In general, the stress on the shift from individual to collective privacy should be welcomed.[263] With regards to collective algorithmic decisions, it would seem that art 22 "does not limit 'profiling' as such to individual profiling but only requires that the decision based on such profiling is addressed to an individual, in a way that has legal or significant effects for him/her as an individual."[264] Therefore, collective profiling is covered by the GDPR when used for individual decisions.

259 Article 29 Working Party (n 9) 26.

260 ibid 26.

261 ibid 26.

262 Kamarinou (n 76) 10.

263 Alessandro Mantelero, 'Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of Data Protection' (2016) 32(2) Computer Law & Security Review 238.

264 Kamarinou (n 76) 10.

IX. Data portability, accountability, and data minimisation

81 Although the focus is on the provisions specifically dedicated to algorithmic decision-making, other rules and principles may affect it. One need only mention data portability, accountability, and data minimisation.

82 The right to data portability could be used to obtain not only information about the logic, significance, and consequences of the algorithmic decision, but also all "the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format."[265] One could use this right to export the profiles used for the algorithmic decision.

83 The principle of accountability, then, may play a positive role. Indeed, in order "to mitigate the risks of automated profiling we must look towards mechanisms that increase the accountability (both through ex ante screening of data mining applications for possible risks and ex post checking of results) and transparency of automated profiling."[266] In particular when relying on the consent-based exception, data controllers will have to document it carefully to prove that consent was explicit.

84 Certain rules should be interpreted broadly, taking into account the characteristics of the phenomenon at hand. For instance, data minimisation and data exclusion, if interpreted narrowly, "may reduce the accuracy of data mining and may deny us the data necessary to detect discrimination in automated profiling."[267] However, the principle of data minimisation means that data should be *adequate*, relevant, and limited to what is *necessary in relation to the purposes* for which they are processed.[268] Arguably, this does not mean that data controllers shall always collect as little data as possible. It means that the quantity must be related to the purpose, provided that the data are adequate. Arguably, the application of artificial intelligence to take decisions that have legal effects can justify the processing of large amounts of data, for at least two interwoven reasons. First, the more data are used to train the algorithm, the more accurate the output may be (big data are 'necessary' for the functioning of artificial intelligence). Second, the processing of a low quantity of data, leading to an inaccurate output, would be 'inadequate' if one has to take a decision with legal consequences (or which similarly significantly affects the individual).

265 GDPR, art 20.

266 Schermer (n 10) 52.

267 ibid 52.

268 GDPR, art 5(1)(c).

X. Algorithmic decisions taken by EU institutions and bodies

85 A brief note, finally, on the algorithmic decision-making carried out by the EU and its institutions and bodies (e.g. e-procurement and e-recruiting). The current rules[269] are more or less the same as the ones laid out in the Data Protection Directive, with the right to be informed about the logic involved in the decision, the right not to be subject to it, and the data controller's obligation to put in place measures to protect the data subject's legitimate interests. The only exception recognised is the express authorisation by national law, EU law, or the European Data Protection Supervisor. In January 2017, the Commission adopted a proposal for a new regulation on the processing of personal data by the EU institutions, bodies, offices, and agencies.[270] The draft provides the same rules as the GDPR as to the information rights (existence, logic, significance, consequences),[271] right to access,[272] right to not to be subject,[273] and mandatory data protection impact assessment.[274]

XI. An overall assessment of the new data protection rules on algorithmic decisions

86 In conclusion, overall the GDPR strengthens the rules on algorithmic decision-making timidly and with some significant flaws, though some positive elements have to be acknowledged. It may well be the case that, as it has been suggested, this regime will act as "legal incentives for technology producers to build accountability mechanisms into the technology."[275] It still holds true that even if Article 15 of the Data Protection Directive and Article 22 of the GDPR show that the promise in terms of providing a counterweight to algorithmic decision-making is tarnished by complexities and ambiguities, they nonetheless shall be regarded as expression of a core data protection principle to be embodied in all data protection instruments.[276]

87 Now, before moving on to the third legal route, one needs to take account of the relation between intellectual property and data protection. It has been shown above that the Software Directive can prevail on the Trade Secrets Directive. It remains to be assessed what happens if there is a clash between trade secrets (and, more generally, intellectual property rights) and the data subject's rights. Under the GDPR, the right of access cannot 'adversely affect the rights and freedoms of others,'[277] which include 'trade secrets or intellectual property and in particular the copyright protecting the software.'[278] However, this provision has been interpreted narrowly by the Article 29 Working Party that observe that intellectual property rights cannot be invoked to deny access or refuse to provide information to the data subject.[279] In allowing the disclosure of an algorithm covered by a trade secret, however, courts shall dictate measures that safeguard the commercial value of the trade secret, for instance by preventing its further disclosure. It is important to note that intellectual property must be balanced with data protection only when it comes to the right of access. Conversely, it is submitted that, in principle, when it comes to the other data subject's rights and data controller's obligations, intellectual property will not be a valid legal basis for exceptions or limitations.

88 Another regime to take into account – and whose interplay with intellectual property and data protection remains partly unsolved – is freedom of information.

269 Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data, arts 13 and 19, recital 29.

270 Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (hereinafter 'draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions'). For the Commission proposal, the first reading Position of the European Parliament and the General Approach of the Council, see Council of the EU 13436/17 of 30 October 2017.

271 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, arts 15(2)(f) and 16(2)(f).

272 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, art 17(1)(h).

273 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, art 23.

274 Draft regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, art 39.

275 Chris Reed et al., 'Responsibility, Autonomy and Accountability: Legal Liability for Machine Learning' (Queen Mary School of Law Legal Studies Research Paper No. 243/2016) 29 <https://ssrn.com/abstract=2853462> accessed 1 March 2018.

276 Bygrave (n 165) 22.

277 GDPR, art 15(4).

278 GDPR, recital 63.

279 Article 29 Working Party (n 9) 17.

E. Freedom of information and access to the algorithm. The Italian panorama

89 In 2015, the French *Commission d'accès aux documents administratifs* obliged the *Direction générale des finances publiques* to release the source code of the computer program used to estimate the income tax of natural persons.[280]More recently, the *TAR Lazio*,[281] administrative court[282] in Italy, stated that an algorithm is a digital administrative act and therefore, under the freedom of information regime, the citizens have the right to access it. This section critically analyses this ruling as a prism to understand the application of the freedom-of-information regime to algorithmic decision-making.

90 Under the Italian Administrative Procedure Act,[283] citizens have the right to view administrative documents and extract a copy thereof, if they have a "direct, specific, and actual interest, corresponding to a legally-protected situation and linked to the document one intends to access."[284] The typical example would be an individual unhappy with the outcome of a public competition (e.g. to become notary public) and, therefore, demands to access the documents relevant to the competition. An important limitation of freedom of information regimes is that they can be actioned only against the State or other public bodies and with regards to administrative documents.[285] The Government and the public bodies can lay out which documents cannot be accessed for a number of purposes listed in the Administrative Procedure Act, including privacy.[286] However, there is case law clarifying that in principle, if the right to access and privacy clash, the former shall prevail, at least in the sense that an access request will not be denied for privacy reasons, but the document may be anonymised.[287] More recently and generally, it has been stressed that freedom of information is a fundamental right and, therefore, the denial to access requests are allowed only in exceptional instances.[288] This approach can also be found in the Privacy Code[289], in which there is a right to access administrative documents even though they contain personal or even sensitive data, because the freedom of information regime "is deemed to be of relevant public interest."[290] The balance is struck slightly differently when it comes to data on health or sexual life. Indeed, the access request will only be accepted if the interest underlying the request is a personality right[291] or other fundamental right or freedom.[292] One may infer that normally the right to access prevails over opposite interests and rights, even in the event the opposite rights were fundamental, unless the computer program implementing the algorithm processes health data or data about the sexual life of the individual. Thus, it is submitted that also the potential clash between freedom of information and intellectual property should normally be resolved in favour of the former. The GDPR will not affect the balance between privacy and freedom of information, since the recently presented draft implementing decree clarified that access to administrative documents and civic access fall outside the scope of the GDPR, at least in the context of its Italian implementation.[293]

91 Only individuals who have a specific, direct, and actual interest in the access to the administrative document can exercise the right of access under the Administrative Procedure Act. However, in 2016, Italy introduced a more general freedom of information regime. Under the Citizen Access Act,[294] the individual has two rights. First, the right

280 Commission d'accès aux documents administratifs, avis 20144578 - 8 January 2015, <http://www.cada.fr/avis-20144578,20144578.html> accessed 1 March 2018.

281 TAR Lazio, chamber III bis, 22 March 2017, No 3769.

282 These courts administer justice mainly when a citizen claim the violation of their legitimate interest by a public body.

283 *Legge* 7 August 1990, No 241 *Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi* (hereinafter 'Administrative Procedure Act'), Articles 22-28. *Decreto del Presidente della Repubblica* 12 April 2006, No 184 *Regolamento recante disciplina in materia di accesso ai documenti amministrativi*.

284 Administrative Procedure Act, art 22(1)(a).

285 Administrative documents are defined in a very broad way, that is "every graphical, photographical, electromagnetic representation (or any other kind of representation) of the content of documents - be they even internal or not related to a specific administrative procedure – which are in the possession of a public body and concern public interest activities, being immaterial the public or private nature of the relevant regime" (Administrative Procedure Act, art 22(1)(d)). For an even broader definition see *Decreto del Presidente della Repubblica* 28 December 2000, No 445 *Testo unico delle disposizioni legislative e regolamentari in materia di documentazione amministrativa*, art 1(1)(b).

286 Administrative Procedure Act, art 24(6)(d). See, for instance, *Regolamento del Comune di Salerno sull'accesso agli atti e sulla tutela della riservatezza dei dati contenuti in archivi e banche dati comunali*, art 5(2)(m).

287 Consiglio di Stato, chamber V, 28 September 2007, No 4999; Consiglio di Stato, chamber VI, 20 April 2006, No 2223; Consiglio di Stato, plenary session, 18 April 2006, No 6.

288 TAR Toscana, chamber I, 10 February 2017, No 200.

289 Decreto legislativo 30 June 2003, No 196, *Codice in materia di protezione dei dati personali* (Privacy Code).

290 Privacy Code, art 59.

291 By personality rights, it is meant rights, such as life and honour, that are absolute and refer to fundamental aspects of the human being. This is a civil law notion, which should not be confused with the common law one, where personality rights are the rights to control the commercial use of one's own name or other aspects of one's identity (name, likeness, etc.).

292 Privacy Code, art 60.

293 Draft GDPR implementing decree, art 55.

294 *Decreto legislativo* 14 March 2013, No 33 *Riordino della*

to access all documents, information, and data (not only administrative documents), if there were an obligation to publish them and the relevant public body infringed it by not publishing.[295] This right (so-called citizen simple access) is absolute and an access request under this provision cannot be denied.[296] Second, a right to access documents that the State or other public bodies are not obliged to publish, justified with the purpose to "favor a generalised control over the pursuit of the institutional functions and over the use of public resources, as well as to promote the participation to the public debate."[297] This citizen generalised access is a limited right.[298] Indeed, the relevant request can be denied for a number of reasons,[299] including data protection[300] and intellectual property.[301]

92 There is another regime that may be used to access algorithms used by the State and other public bodies, even though its scope is very narrow. As of 14 September 2016, under the Public Administration Code,[302] legal and physical persons have the right to reuse computer programs and other "solutions" in order to "adapt them to their needs".[303] Therefore, the State or other requested public body have an obligation to make the relevant source code publicly available "alongside the documentation"[304] under a free and open-source license. However, the requested body can deny access in three scenarios if the computer program or the solution owned by the State or public body were not developed "based on the specific indications by the public customer."[305] The denial may be justified also by *ordre public*, national security, defence, and elections.[306]

93 Let us focus on the recent case that applied the Administrative Procedure Act in order to recognise the right to access the source code of the computer program implementing the algorithm used by the Ministry of Education, University and Research with regards to the mobility of the teaching staff; the algorithm had been commissioned to a private company (HPE Services s.r.l.). The teachers' trade union claimed that they could not defend their members' right with regards to the mobility procedures if they were not allowed to access the algorithm. The computer program was used to manage the transfer of the teaching staff between provinces and the outcome of the procedure was solely determined by the algorithm. This means that, should the requirements be met (personal data, decision with legal effect, etc.), the applicant may exercise the rights recognised by the GDPR with regards to algorithmic-decision making.[307]

94 In the case at hand, the applicant sought to exercise the right to access under the freedom of information regime. However, this was denied by the Ministry of Education for a number of reasons. Firstly, the source code was not an administrative document (and the right to access under freedom of information can be exercised only with administrative documents).[308] Secondly, the computer program was covered by copyright. The court, however, dismissed both arguments.

disciplina riguardante il diritto di accesso civico e gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni (Citizen Access Act), as amended by the *Decreto legislativo* 25 May 2016, No 97 *Revisione e semplificazione delle disposizioni in materia di prevenzione della corruzione, pubblicità e trasparenza, correttivo della legge 6 novembre 2012, n. 190 e del decreto legislativo 14 marzo 2013, n. 33, ai sensi dell'articolo 7 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche* (Prevention of Corruption Act).

295 Citizen Access Act, art 5(1).

296 Unless the public body proves that there was no obligation to publish or that the document, information or data is already published.

297 Citizen Access Act, art 5(2).

298 Most European jurisdictions have similar provisions. In the UK, the Freedom of Information Act 2000, that covers all recorded information held by a public authority (Information Commissioner's Office, *Freedom of Information Act Awareness Guidance No. 12*). However, an access may be denied for a number of reasons, including trade secrets and other commercial interests (Freedom of Information Act 2000, Section 43). It is notable that, unlike other commercial interests, if the algorithm is covered by a trade secret, the access request may be denied without considering whether or not the release may cause harm (Information Commissioner's Office, *Freedom of Information Act Awareness Guidance No. 5*).

299 Citizen Access Act, art 5 *bis.*

300 Citizen Access Act, art 5 *bis* (2)(a).

301 Citizen Access Act, art 5 *bis* (2)(c).

302 *Decreto legislativo* 7 March 2005, No 82 *Codice dell'amministrazione digitale* (Digital Administration Code), as amended by *Decreto legislativo* 26 August 2016, No 179 *recante "Modifiche e integrazioni al. Codice dell'amministrazione digitale, di cui al decreto legislativo 7 marzo 2005, n. 82, ai sensi dell'articolo 1 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche".*

303 Digital Administration Code, art 69(1).

304 Digital Administration Code, art 69(1). The wording is very vague; it is likely to refer primarily to all the documentation necessary to adapt the computer program to the applicant's needs.

305 Digital Administration Code, art 69(1).

306 In the UK, there is the right to access datasets for reuse and it is broader than the Italian regime, because it regards all copyright works (Freedom of Information Act 2000, s 11A).

307 Currently, in Italy, the Privacy Code does not regulate algorithmic decision-making. The GDPR, being a regulation as opposed to a directive, will play an important role in strongly harmonising the relevant national regimes, in some instances by innovatively recognising the right to be informed about and object to algorithmic decision-making (e.g. in Italy), in others by updating the existing regime (e.g. in the UK). As seen above, the implementing measures of said countries seem to partly or completely overlook the matter.

308 Administrative Procedure Act, art 22(1)(d).

95 Given that, with the current development of AI and kindred technologies, public bodies can increasingly replace human procedures with algorithmic ones, the court held that the use of the algorithm cannot act as justification for restricting the scope of application of the freedom of information regime. Let us imagine what would happen if all procedures were handled by algorithms and the freedom-of-information requests were not applicable to algorithmic documents: the said regime would still exist in the books, but no longer in practice.

96 The conceptual first step is recognising the existence of the concept of a digital administrative document. In a digital administrative document, an algorithm replaces a human agent acting on behalf of a public body; this is allowed only with regards to the non-discretionary administrative activities.[309] Indeed, non-discretionary power is compatible with the way computer programs work, because the latter can translate facts and legal data into code, thus bringing to an immutable conclusion through formalised reasoning.[310] This passage of the ruling reinforces this paper's argument that algorithms cannot replace human judges (and other decision-makers) because interpretation is ubiquitous and it is an intrinsically discretionary process.

97 This said, the court needed to qualify the computer program itself as a digital administrative document, otherwise no access to the source code could be granted (at least under this regime). The computer program qualifies as a digital administrative document because it materialises the ultimate will of the public body in a way that is able to create, modify, or extinguish the individual's legal positions. Consistently with the technology neutrality principle, the relevant statutory provision describes the 'administrative document' in a very broad way by encompassing also the electromagnetic representation of a document and any other form of representation.[311] Therefore, there is no problem in considering a computer program implementing an algorithm as an administrative document (if the other legal requirements are also met).[312] It may be conceded that, strictly speaking, a computer program is not a document in itself. However, recognising the right to access only to the final document resulting from the algorithmic procedure would equal denying the access request, because without the source code it may prove hard to understand the rationale of the final decision. The right to access often serve the purpose of lodging a complaint against a public body if the final decision affected the individual's rights or legitimate interests. However, it is unlikely that such a claim would be successful, if the individual does not have access to the rationale of the final decision (which means also accessing the source code, if the decision is algorithmic). Indeed, it is believed that a narrow interpretation of an 'administrative document' would not comply with the right to an effective remedy and to a fair trial as enshrined in the Charter of Fundamental Rights of the EU[313] and in the European Convention of Human Rights.[314]

98 One may object that granting the access in this case would be tantamount to granting access to the source code of the computer program (e.g. Microsoft Word) used to write an administrative document. Such an argument would be based on a wrong understanding of what is a digital administrative document. Indeed, the court distinguishes between documents drafted with the aid of a computer and electronically programmed documents, where the software finds and links data and norms. The latter is a digital administrative document (the source code of which is accessible) because it constitutes the final decision; it is not a mere aid to draft it.[315] This paper joins those who underline that "the electronic processing *is* the document, it represents it, it makes it known externally, it becomes the form of the document, thus being legally relevant *in its electronic form*, regardless of its paper transcription."[316] The

309 In Italy, the discretional power of the public administration is a fundamental principle, whereas only in a limited number of instances the State or other public body take a non-discretionary decision (with the content as well as the requirements rigidly predetermined by the law), for instance when an authorisation shall be released as a necessary consequence of the positive assessment of the existence of certain requirements. Some authors affirm that administrative power is always discretionary (e.g. Fabio Massimo Nicosia, *Potere ed eccesso di potere nell'attività amministrativa non discrezionale* (Jovene 1991), but this theory is not widely accepted (e.g. Paola Rossi, *Il riesame degli atti di accertamento* (Giuffrè 2008)).

310 The Italian Court of Cassation defined the digital administrative document in a narrow way by including only those documents which are directly and automatically processed from the computer, in as much as they do not require discretional assessments and argumentations linked to the specificities of the case at hand (Corte di Cassazione, chamber I, 28 December 2000, No 16204).

311 Administrative Procedure Act, art 22(1)(d).

312 In particular, the document must be in a public body's possession and it must regard public interest activities (Administrative Procedure Act, art 22(1)(d)). It is immaterial if the algorithm was developed as a consequence of contract (a private law tool), as long as the relevant activity is of public interest, which is the case here, given that the purpose of the program is to improve the management of a public service (education).

313 Art 47.

314 Art 6, art 13.

315 Contrary to what was held by the court, some scholars affirm that only the administrative document drafted with the aid of a computer is a digital administrative act. See Alfonso Contaldo and Luigi Marotta, 'L'informatizzazione dell'atto amministrativo: cenni sulle problematiche in campo' (2002) 18(3) Diritto dell'informazione e dell'informatica 576.

316 Massimiliano Minerva, 'L'attività amministrativa in forma elettronica' (1997) 4 Foro amministrativo 1300, italics

very broad definition of administrative document is seen by the court and by legal scholars as a shift from a focus on the pedigree of the document, to its function:[317] if the function is administrative (as in concerning the public interest), then it is immaterial how the document was formed and access shall be granted in any event, if the general requirements are met. This said, it is important to stress that the court stated that electronically programmed documents are not allowed when it comes to the exercise of discretionary power,[318] due to the difficulty "which is scientific as opposed to legal, to map the reasoning underlying the document,"[319] if this is the outcome of an algorithmic procedure (and not simply drafted by a human being with the aid of a word processor). Again, there is no place for algorithmic decisions where the relevant process is discretionary.[320]

99 After recognising the right to access the computer program, the court went on to state that providing the applicant with the mere description of the algorithm and of its functioning is not a sufficient response.[321] Only the access to the source code is. Indeed, the Ministry of Education had responded to the access request by describing the algorithmic procedure (collection of input data, appointment to a certain school, distribution of the results), as well as reporting some case studies. The court, however, states, "the assessment of the functionality of the algorithm or of programming errors can be carried out exclusively in light of the knowledge"[322] of the source code. This should be accompanied by a thorough explanation of the rationale and of the consequences of the decisions, especially if personal data is involved.

100 Finally, as to the clash with the copyright on the computer program, the steps to follow are: i. Assessment of copyright subsistence; ii. Authorship and ownership; iii. Infringement; iv. Exceptions.

101 The subsistence, authorship, and ownership of the copyright do not seem to be problematic.[323] Even though there is no evidence on the point, the court assumes that the Ministry of Education owns the program under a license with HPE Services s.r.l., which retains authorship and the moral rights.[324]

102 The court goes on to observe that the purpose of the access does not conflict with the economic interest protected by copyright.[325] On this point, the court is not clear as to whether it is dealing with the assessment of infringement or with the exceptions. In the latter event, this would be a peculiar ruling, because it would take a flexible "fair use"[326]-like approach to copyright exceptions,

added.

317 Carmelo Giurdanella and Elio Guarnaccia, *Elementi di diritto amministrativo elettronico* (Halley 2005) 24.

318 Most scholars agree, see e.g. Contaldo (n 312) 580.

319 TAR Lazio, chamber III bis, 22 March 2017, No 3769. This idea was first expressed by A Ravalli, 'Atti amministrativi emanati mediante sistemi informatici: problematiche relative alla tutela giurisdizionale' (1989) 2 Trib. Amm. Reg. 261. The traditional theory that presents a dichotomy discretionary-non-discretionary and allows algorithmic decisions (or electronically processed administrative documents) only with regards to the latter is open to criticism. However, this is not because, as Ravalli thinks, even discretionary administrative activities are rational logical processes based on predetermined criteria (which is debatable). The point is that interpretation is always discretionary and even non-discretionary power is exercised through interpretation (given that the dichotomy interpretation-application is untenable, as shown by Hart; see Viola (n 29) 50).

320 This passage may be interpreted as the court espousing that line of thought whereby the admissibility of algorithmic decisions (or electronically processed administrative documents) depends not on the nature of the power, but to the scientific possibility to map the reasoning underlying the document (Giurdanella (n 314) 32; Michele Corradino, 'Inquadramento generale dell'atto amministrativo elettronico' (Convegno DAE 2004). However, before referring to the importance of the said scientific possibility (or the lack thereof), the court is adamant in reaffirming the old contraposition. Indeed, the court states that "we can easily agree that administrative documents which are the output of an algorithmic procedure are admissible with regards to the non-discretionary activity of the public bodies" (TAR Lazio, chamber III bis, 22 March 2017, No 3769). It then goes on to observe that "it is evident that different considerations apply to the discretionary activities" (ibid.). The reference to the fact that the admissibility of this kind of digital administrative act does not depend on the qualification of the activity as discretionary (but it would depend on the possibility of mapping the underlying reasoning) is introduced by a dubitative form ("it may be possible to assume that") and it seems an obiter dictum. One may infer this by the observation that "we believe that we can disregard the exam of this very interesting legal question" (ibid.).

321 Some believe that the description of the algorithm could solve the problem of making the citizen understand the software used by the public body. See Daniele Marongiu, 'Gli atti amministrativi ad elaborazione elettronica: la compilazione di un "pre-software" in lingua italiana' (Quaderni del DAE 2003) <http://www.cesda.it/quadernidae/pdf/MARONGIU_DAE2003.pdf> accessed 1 March 2018.

322 TAR Lazio, chamber III bis, 22 March 2017, No 3769.

323 The court accepts the Ministry of Education's allegations on the point, because there are no elements that may suggest that there is no copyright on the computer program at hand.

324 Transactions regarding moral rights (e.g. paternity waivers) are not enforceable under Italian copyright law (*Legge* 22 April 1941, No 633 *Legge a protezione del diritto d'autore e di altri diritti connessi al suo esercizio* (Copyright Act), art 22).

325 In Italy, the author of a copyright work has the exclusive right to use the work for economic purposes (Copyright Act, art 12(2)).

326 This is the doctrine of copyright exceptions in the US. It does not revolve around a list of permitted uses, but it is a flexible principle that enables the judge to assess all the

usually interpreted by applying the so-called three-step test, revolving around an exhaustive list of permitted uses.[327] There is currently no copyright exception for non-commercial use or for purposes of freedom of information. The access to the source code for this purpose may not conflict with the normal exploitation of the work and may not prejudice the interests of the author. However, the third step requires that the exception be expressly provided by the law, which currently does include a general exception for non-commercial acts. Conversely, the point should be better construed as meaning that there can be no infringement because the restricted act is not the distribution of the copyright work, but its distribution for commercial purposes. Indeed, the heading of the chapter of the Copyright Act on the restricted acts is "Protection of the economic use of the work"[328] and the first relevant provision recognises the "exclusive right to economically use the work within the limits of the Act."[329] From this perspective, the clash between freedom of information and copyright is merely ostensible, because the right to access administrative documents does not interfere with the uses of computer programs that are restricted by the law. Additionally, a different conclusion would have led to an unacceptable difference of treatment depending on the technological solution adopted. It is obvious that, in principle, public bodies own copyright on the documents they produce. However, it would be absurd to claim that a freedom of information request can be denied because the public body owns the relevant copyright. This would equal sterilising the right to access. Accordingly, the discretional adoption of a more modern technology cannot justify different considerations. Therefore, just like copyright could never be the basis of an access denial under the analysed regime, it will never justify the access denial with regards to computer programs.

103 An argument of the Ministry of Education was, then, that the so-called citizen generalised access request can be denied if necessary to avoid an actual prejudice to intellectual property.[330] However, the right to access under the Administrative Procedure Act (which is the one relevant here) and the citizen access are entirely different things. Their purposes are discrete. The former does not encompass a right to a generalised control over the public bodies:[331] it serves the purpose to enable the individuals to defend their rights and interests which may be affected by an administrative document. This generalised control, conversely, is the purpose of the citizen access rights under the Citizen Access Act. The requirements of the right to access administrative documents and the citizen access rights (both simple and generalised) are different; therefore, all the remedies can operate in parallel. The balance will have to be struck differently. On the one hand, the former requires access to more detailed information, because it serves the purpose of preparing a claim. On the other hand, under a citizen access regime, even less granular information will be sufficient (e.g., the description of the algorithm may suffice under this regime). The court states that, therefore, it may be that whereas a citizen access is denied, it may be accepted with regards to the same document if the same individual exercises the right to access administrative documents.

104 It is submitted that the court may have brought into play three more considerations. First and foremost, *ubi lex voluit dixit, ubi noluit tacuit*. The lawmaker expressly accepts that an access request can be denied for intellectual property purposes under the citizen access regime. However, the fact that the legislator does not provide a similar exception with regards to the right to access administrative documents constitutes evidence of the untenability of an intellectual property exception to the said right. Second, intellectual property is mentioned in the citizen access regime as an example of "economic and commercial interests."[332] Therefore, since it has already been proven that the access to the source code would not conflict with the use of the program for commercial purposes, even if the exception were extended to the right to access administrative documents, it would not apply in the case at hand. Third, the exceptions to the citizen access are allowed only if "necessary to avoid an actual prejudice" to the listed interests (including intellectual property). Arguably, denying access to the source code may not always be necessary to avoid such prejudice (for instance, if the applicant agrees to make a non-commercial use of it). Given that there is no intellectual property exception to the right to access administrative documents, one should bear in mind that also trade secrets and patents might not be used to prevent the said access. This is particularly important from our perspective, given the pivotal

circumstances of the case to ascertain whether the use of a copyright work was fair.

327 The exception must fall within the exhaustive list of the Copyright Directive (Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society), not conflict with the normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author (art 5(5)).

328 Copyright Act, *Titolo* I, *Capo* III, *Sezione* I.

329 Copyright Act, art 12(2). See arts 64 *bis* – 64 *quater* for the specific provisions on computer programs (the general rule on the economic use, however, applies also to computer programs).

330 Citizen Access Act, Article 5 *bis* (2)(c).

331 Administrative Procedure Act, Article 24(3). See, for instance, Consiglio di Stato, chamber V, 25 September 2006, No 5636.

332 Citizen Access Act, art 5 *bis* (2)(c).

role of trade secrets in keeping algorithms opaque.

105 As a consequence of the lack of the elements of infringement, of the inexistence of an intellectual property exception to the right to access administrative documents, as well as of the general assertion whereby "the nature of copyright work does not represent a justification for access denial,"[333] the court recognises the right to access the source code, provided that the applicant uses the information exclusively for the purposes that legitimised the claim (the right of the teachers' trade union to defend its members' rights).

106 For all the reasons analysed above, the court found in favour of the teachers' trade union and, therefore, annulled the access denial and ordered the Ministry of Education the release of a copy of the source code of the computer program implementing the algorithm used by the Ministry in handling the teachers' mobility.

107 The right to access administrative documents may be seen as a weak tool when it comes to the transparency of the algorithmic decisions taken by the State and other public bodies. Indeed, especially in AI / black box contexts, accessing the source code of the computer program implementing an algorithm does not provide the applicant with valuable and / or intelligible information.[334] However, denying such access would conflict with the fundamental right to an effective remedy, because an individual could hardly be successful in a claim against a public body, if they cannot access the rationale of an algorithmic decision affecting their rights and legitimate interests.

108 Some scholars suggest that, in the future, artificial intelligence will be used to adopt algorithmic administrative documents even when it comes to discretionary activity, with the possibility of leaving room for the human intervention in the most difficult cases.[335] They maintain that this is only a prediction but given the current developments of natural language processing and machine learning, arguably the relevant tools are already available. Even though it cannot be said that artificial intelligence should be banned altogether when it comes to discretionary power, it is believed that some room for *ex-ante* human intervention should always be left for a number of reasons, including the fact that all administrative activities (like all interpretive operations) are to some extent discretionary. This does not mean, however, that citizens cannot exercise the right to access under the freedom of information regime if the relevant administrative activity is non-discretionary. It means that public bodies are not allowed to use AI when they are exercising a discretionary power.

109 The question remains as to what citizens can do if public bodies start taking decisions against them even in the discretionary realm. The remedy described in this section operates *ex post*, once the decision has already been taken. Similarly, the copyright and patent exceptions may constitute a useful *ex-post* tool, but their scope is quite limited. From an *ex-ante* perspective, however, it may be argued that a potentially affected individual could obtain an injunction to prevent a public body from taking an algorithmic decision by using the data protection tool under Article 22 of the GDPR. Therefore, an integrated approach to the remedies against algorithmic decisions should be taken.

F. Conclusions

110 This study presented ten arguments against algorithmic decision-making, as well as three routes available to those affected by algorithms. As pointed out by some scholars,[336] the most important thing is providing individuals with the means to challenge adverse algorithmic decisions. To do so, intellectual property, data protection, and freedom of information provide adequate protections, particularly if one takes an integrated approach. National implementations of the GDPR should be a precious opportunity to detail the procedures to challenge algorithmic decisions, even though it does not seem that this is the direction that is being taken.

111 Intellectual property enables the legitimate user of a software implementing an algorithm or of an algorithm-related patent to carry out certain acts (study, observation, etc.) without the intellectual property owner's consent. Whilst these quasi-rights allow the user to try and understand the algorithm by themselves, they do not give them a positive right to demand the intellectual property owner's cooperation.

333 TAR Lazio, chamber III bis, 22 March 2017, No 3769.

334 It is not a coincidence that the applicant is not an individual, but a trade union, which is likely to have the resources to make sense of a source code. The fact that a lay person could hardly understand a source code has been used as an argument against the recognition of computer programs as digital administrative documents. However, the court points out that the choice of an innovative tool cannot deprive the citizens of the right to access administrative documents and that, anyway, the applicant may avail themselves of the collaboration of an IT person to decipher the code.

335 Giurdanella (n 314) 33, referring to Giovanni Sartor, *Le applicazioni giuridiche dell'intelligenza artificiale* (Giuffrè 1990) and Giovanni Sartor, 'Gli agenti software: nuovi soggetti del cyberdiritto?' (2002) Contratto e impresa 465.

336 Keats (n 4) 1.

112 Conversely, a freedom of information request allows all citizens to impose upon public bodies, under certain circumstances, an obligation to release the source code of computer programs that implement algorithms, while explaining the logic involved in the relevant decision. The main shortcoming of this regime is the limitation to public defendants. Much will depend on how courts will strike a balance between freedom of information and intellectual property. In Italy, the former prevails. In turn, arguably, the UK tend to favour the interests of the intellectual property owners.

113 The only *ad-hoc* regime against algorithmic decisions is provided by art 22 of the GDPR. One may criticise some aspects of this provision. For instance, it applies only to decisions "*solely* based on automated processing" means. This paper's suggestion is to recognise the right not to be subject to an algorithmic decision every time that there is not a human being taking the final decision substantially, as opposed to formally. In spite of its shortcomings, art 22 is clear and detailed in laying out the general principle that businesses, governments, judges, and other data controllers should not make decisions based solely on algorithmic processes. Under certain circumstances (e.g. explicit consent), such decisions can be made, but informing the data subject and allowing him or her to access to the logic involved in the decision, its significance, and the envisaged consequences. Much will depend on the national implementing measures. The UK Data Protection Bill risks not ensuring compliance with the GDPR, thus exposing the UK to the possibility of being considered as 'inadequate' in the context of cross-border EU-UK data transfers.

114 It is submitted that only a document which includes *both* the algorithm used and an explanation of the logic and consequences in non-technical terms would comply with the GDPR as interpreted in light of the Charter of Fundamental Rights of the EU and the European Convention on Human Rights. Then, the right to a human judge is paramount, because the right to access and to be informed may prove useless. Indeed, when artificial intelligence is used, it is sometimes unfeasible to access the relevant rationale. To the legal black box created by intellectual property rights, one needs to add the technical black box and the organisational one.

115 Practically, if the algorithmic decision is based on personal data, this latter route is preferable. If not and the decision-maker is a public body, one should opt for a freedom of information request. If a private decision-maker (e.g. a bank) makes an algorithmic decision based on non-personal data, then the route will be that of intellectual property exceptions. The freedom-of-information remedies operate *ex post*, once the decision has already been taken. In turn, the copyright and patent exceptions may be used before any decision is made, but only to access the algorithm, not to prevent the decision-maker from proceeding algorithmically. The only regime that prevents algorithmic decisions is the one provided by the GDPR.

116 The trust in artificial intelligence and algorithms derives from the belief that non-human agents are unbiased, and their decisions are not affected by passions and ideologies. In fact, algorithms are as biased as the people who trained them, but in a less transparent and accountable way. The more important algorithms will become, the more we will want them to embed our values (and, therefore, our ideologies and biases).[337] Further research should be carried out by diverse (also in terms of gender, ethnicity, etc.) multidisciplinary teams in order to find solutions to open the technical, organisation, and legal black boxes and to ensure fair algorithmic decision-making. Indeed, only a strong humanist stance will be able to reduce algorithmic bias.

117 This paper is a humanist manifesto. It is, indeed, permeated with the belief that we should trust our fellow human beings over the algorithms, despite developments in artificial intelligence allowing the deployment of increasingly refined legal applications. This does not mean that we should reject the use of algorithms altogether. For instance, judges shall use them to improve the quality and consistency of their decisions. However, they shall not let algorithms decide in their stead. In order to better understand how to make the human-algorithm cooperation work best, it has become crucial to shift the focus from the definition of algorithms, artificial intelligence etc. to the understanding of what makes us human.[338]

Acknowledgements

The author is grateful to Sue Farran, Paul Dargue, and Tony Ward for comments on previous drafts of this article. This work has greatly benefited also from the feedback received at the «XXVIII World Congress of Philosophy of Law and Social Philosophy» (Lisbon, 17 July 2017), at the «Café & Chat: Quem Governa os Algoritmos?» (IRIS - Instituto de Referência em Internet e Sociedade

337 The trends of data protection by design, ethics by design, etc. may be explained as a tendency to anthropomorphise non-human agents. See, for example, British Standard Institution, *Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems* (BSI 2016); *cf* Guido Noto La Diega, 'The European strategy on robotics and artificial intelligence: too much ethics, too little security' (2017) 3(2) European Cybersecurity Journal 6.

338 For a positive step in this direction, see Bett M Frischmann and Evan Selinger, *Re-engineering humanity* (CUP 2018).

and GNET – UFMG, Faculdade de Direito da Universidade Federal de Minas Gerais, Belo Horizonte, 18 August 2017), at the research seminar organised by NINSO The Northumbria Internet & Society Research Interest Group (Newcastle upon Tyne, 8 December 2017), and at a guest lecture given at the Schmalkalden University of Applied Sciences (Schmalkalden, 21 December 2017). Thanks to Katie Atkinson and Giulia Caffarelli for the insight into the AI debate. The author's appreciation goes to the anonymous referees for the helpful comments and to Philipp Schmechel for the patient editing. Views and errors are solely the author's responsibility.